Single cell analysis reveals all-or-none G1 arrest decisions upon TGFβ stimulation

Dissertation

zur Erlangung des akademischen Grades
doctor rerum naturalium
(Dr. *rer. nat.)*
im Fach Biophysik

eingereicht an der
Lebenswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin
von

MSc. Guoyu Wu

Präsidentin der Humboldt-Universität zu Berlin
Prof. Dr.-Ing. Dr. Sabine Kunst
Dekan der Lebenswissenschaftlichen Fakultät
Prof. Dr. Bernhard Grimm

Gutachter/innen: Dr. Zhike Zi
Dr. Jana Wolf
Prof. Dr. Dr. h. c. Edda Klipp

Tag der mündlichen Prüfung: 17. 04. 2019

Abstract

The transforming growth factor-β (TGFβ) exerts diverse effects on regulating numerous biological processes. Especially, the cytostatic effect of TGFβ is important for maintaining tissue homeostasis and preventing proliferative disorders, like cancer. Previous studies on the regulation of cell cycle by TGFβ were conducted at the population level, and often through physical or chemical synchronization, which neglected cellular heterogeneity and might introduce artifacts.

To understand how individual cells decode and integrate TGFβ signals into cell proliferation decisions, we quantitatively characterized both TGFβ signaling dynamics and cell cycle progression in asynchronous cells by live cell imaging. Combining experimental and theoretical studies, we demonstrated that TGFβ triggers all-or-none G1 arrest, which is both dose-dependent and phase-dependent. When exposed to TGFβ during S/G2/M phase, cells undergo an inherited, delayed arrest at the next G1 phase. In addition, after one pulse of TGFβ stimulation, cells are refractory to further TGFβ treatments. Considering the importance of single cell information and challenges in automatic cell tracking, we proposed a Population to Single cell framework (P2S framework) to infer single-cell lineages from population dynamics.

Taken together, this work provides new insight into strategies to control cell proliferation by manipulating TGFβ signaling.

Key words: Cell cycle control, TGFβ signaling, heterogeneity, Multi-scale modeling

Zusammenfassung

Der transformierende Wachstumsfaktor-β (TGFβ) übt verschiedene Wirkungen auf die Regulierung zahlreicher biologischer Prozesse aus. Insbesondere die zytostatische Wirkung von TGFβ ist wichtig, um die Homöostase in Geweben aufrechtzuerhalten und proliferative Störungen, wie in Krebs, zu verhindern. Frühere Studien zur Regulation des Zellzyklus mit TGFβ wurden auf Populationsebene, oft durch physikalische oder chemische Synchronisation durchgeführt. Dabei wird die Heterogenität auf zellulärer Ebene vernachlässigt und die Anfälligkeit gegen potenzielle Artefakte erhöht. Um zu verstehen, wie einzelne Zellen TGFβ-Signale entschlüsseln und diese in die Entscheidung zur Zellproliferation integrieren, wurden sowohl die Dynamik der TGFβ-Signale als auch die Zellzyklusprogression in asynchronen Zellen durch „Live Cell Imaging“ quantifiziert. In Kombination von experimentellen und theoretischen Studien wurde gezeigt, dass TGFβ einen „Alles-oder-Nichts-G1- Stillstand“ auslöst, der sowohl dosisabhängig als auch phasenabhängig ist. Wenn die Zellen während der S / G2 / M-Phase TGFβ ausgesetzt werden, erfahren sie in der darauf folgenden G1-Phase einen ererbten, verzögerten Stillstand. Zusätzlich sind die Zellen nach einem TGFβ-Stimulationsimpuls für weitere TGFβ-Behandlungen unempfindlich. In Anbetracht der Bedeutung von Einzelzellinformationen und den Herausforderungen bei der automatischen Zellverfolgung wurde ein Rahmenkonzept von „Population to Single Cell“ (P2S-Framework) erarbeitet, um von der Populationsdynamik auf die Abstammung einzelner Zellen zu schließen. Zusammengefasst bietet diese Arbeit neue Einblicke in Strategien zur Kontrolle der Zellproliferation durch Manipulation der TGFβ-Signalgebung.

Schlagwörter: Zellzykluskontrolle, TGFβ-Signalgebung, Heterogenität, Multi-Scale-Modellierung

Contents

Chapter 1

Introduction

Escaping the growth inhibition of cytokine is one of the hallmarks of cancer[1]. In normal tissues, the homeostasis is tightly controlled by several signaling cues, such as EGFR[2], FGF[3], Wnt[4] and TGFβ superfamily[5] signaling. The agonist and antagonistic cytokines are active in balance to maintain proper cell proliferation and cell death[6,7].

Transforming growth factor-β (TGFβ) is a pleiotropic cytokine and elicits context-dependent responses. Numerous cellular processes are controlled by TGFβ, which include cell proliferation, differentiation, migration and apoptosis[8]. Especially, the growth inhibitory effect, which makes TGFβ to be a tumor suppressor, has been studied intensively. TGFβ strongly inhibits proliferation by inducing cell cycle arrest at G1 phase in most cell types, including epithelial, endothelial and immune cells. Dysfunction of TGFβ signaling leads to hyper-proliferation diseases, like cancer[9]. In the past decades, TGFβ signaling pathway has got an explosion of interest as a potential clinical drug target[10].

In this chapter, we firstly give a brief overview on TGFβ signaling and its cytostatic function. Then we introduce the methods of single cell analysis on signaling events and how we can take advantage of mathematical modeling to better understand cellular behaviors.

1.1 The TGFβ signaling

Since 1978, when TGFβ was discovered, tremendous efforts have been put into the identification and characterization of TGFβ signaling[11]. The main components in the pathway and molecular mechanisms have been studied extensively.

1.1.1 The canonical TGFβ/Smad signaling pathway

In the canonical TGFβ signaling network, the binding of TGFβ ligand to the TGFβ type II receptor (TβRII) initiates the signaling cascade (Figure 1.1). After that, the TGFβ type I receptor (TβRI) is recruited to the complex, being phosphorylated in the GS domain. Then the activated TβRI specifically recognizes the receptor Smads (R-Smads, including Smad2 and Smad3) and phosphorylates their C-terminal serine residues[12].

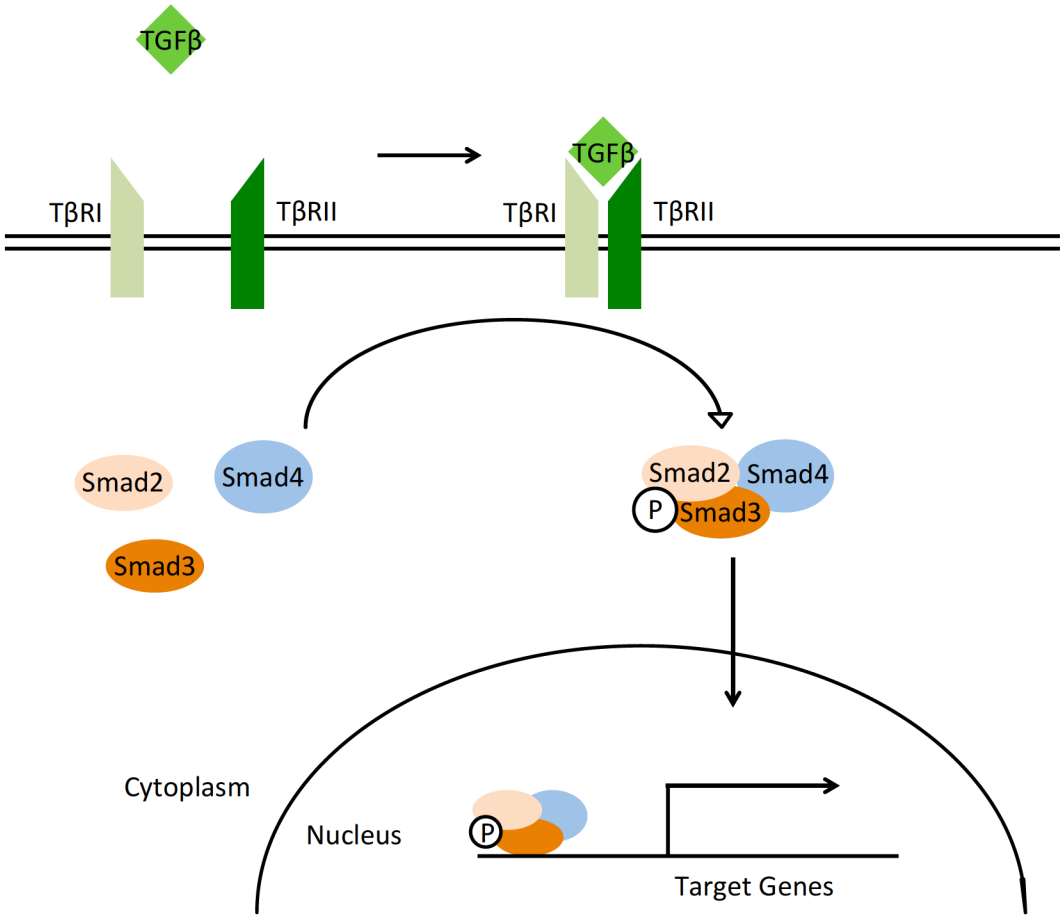


Figure 1.1 Diagram of canonical TGFβ signaling pathway. The signaling pathway starts from the binding of TGFβ to TβRII. Then TβRI is recruited to the complex and phosphorylated. After that, receptor Smads (R-Smads, including Smad2 and Smad3) translocate to the TβRI-TβRII complex and are phosphorylated. Then the phosphorylated R-Smads form oligomers with Smad4 and translocate to the nucleus, and then modulate the transcription of target genes.

Smads continuously shuttle between cytoplasm and nucleus and the shuttling is mediated by several importins and exportins[13]. When the R-Smads are phosphorylated and form a complex with Smad4, TAZ binds to the Smad complex and retains it in the nucleus[14]. Thus the nuclear exporting rates of Smads decrease and a TGFβ-triggered nuclear Smads accumulation is observed[15].

In the nucleus, the Smads protein complex incorporates with DNA-binding cofactors, including co-activators (such as p300, SMIF and CBP[16]) or co-repressors (such as SNON, SIP1 and EVI1[17]), and selectively binds to target promoters, modulating the transcription of target genes.

Hundreds of genes are activated or repressed by TGFβ. Several subsets of genes are identified to respond to TGFβ[9,18](Table 1-1). These genes have been categorized into functional groups, which are further divided into up-regulated/down-regulated subtypes. They regulate the interactions of the cells with the environment, by controlling the extracellular matrix components or cytokines production, or by modifying the transcriptional regulators and signaling networks[9].

Table 1-1 Genes in response to TGFβ shared by the breast, skin and lung epithelial cells

Functional group	TGFβ response	Genes
Cytostatic program	Up-regulated	*CDKN2A, CDKN2B*
	Down-regulated	*ID1, ID2, ID3, c-MYC*
Paracrine network	Up-regulated	*IL11, JAG1, VEGF, CTGF, FSTL3, ANGPT4*
	Down-regulated	*BMP4, IL1β*
Signaling network	Up-regulated	*BMPRII, RHOGEF114, VDR, EPHB2, MEKK4*
	Down-regulated	*PGE-R4, LDLR, βAR-2*
Transcriptional network	Up-regulated	*JUNB, ETS2, c-JUN, ATF3, PIM1, GADD45B, MAD2, MAD4*
	Down-regulated	*TRIP-Br2, C/EBPδ, MRG1*
Extracellular matrix	Up-regulated	*PAI1, ADAM19, uPA, Col VI-A1, ITGα5, ITGβ6*
Negative feedback	Up-regulated	*SMAD7, SMURF1, SMURF2, SNON*
Other response	Up-regulated	*TBX3, MN1, IGl, SIAT4A*
	Down-regulated	*SPRY2, IAP3*

* The table is adapted from Peter M. Siegel et al (2003).

Smads undergo continuous nucleocytoplasmic shuttling during active TGFβ signaling, with cycles of phosphorylation in cytoplasm and dephosphorylation in the nucleus. Nuclear accumulation of Smad proteins requires continuous receptor activity, and it is crucial to TGFβ-induced transcription. The duration of nuclear Smads accumulation reflects the duration of active TGFβ signaling[19], which is important in determining the specific biological responses[20].

The TGFβ/Smad canonical signaling plays an essential role in most TGFβ-induced responses. Several non-canonical signaling pathways are also involved in TGFβ signaling, which are Smad-independent.

1.1.2 The non-canonical TGFβ signaling pathway

Besides Smads, several other factors can be activated by TGFβ, such as c-Jun N-terminal kinase (JNK), phosphatidylinositide 3-kinase (PI3K)/AKT, TNF receptor-associated factor 4/6 (TRAF 4/6), Rho-like GTPases (Rho) and mitogen-activated protein kinases (MAPKs)[21] (Figure 1.2).

The activation of non-Smad pathways results in diverse cellular responses. Among them, Ras/Erk and RhoA signaling are essential for epithelial to mesenchymal transition (EMT), which is one of the fundamental biological functions of TGFβ. During EMT, the adherent junctions and affiliated proteins, e.g. E-cadherin, are down-regulated and the actin stress fibers are induced, by which the cell acquires motile and invasive properties[22]. In the advanced stages of tumorigenesis, the TGFβ-induced EMT promotes tumor metastasis.

JNK/p38 MAPK signaling pathway may be the best characterized non-canonical TGFβ signaling pathway. In addition to the canonical Smad signaling pathway, the TRAF6-TAK1-JNK/p38 signaling pathway also plays an essential role in TGFβ-induced apoptosis, which is recognized as the tumor-suppression function of TGFβ. Interestingly, the TRAF6-TAK1-JNK/p38 signaling pathway is also crucial in TGFβ-induced EMT, which promotes tumor metastasis in tumorigenesis. The TRAF6-TAK1-JNK/p38 signaling cascade is initiated by the binding of TGFβ ligand to the TGFβ type I/II receptors. The activated complex interacts with TRAF6 and activates the lysine-63 (K63)-linked poly-ubiquitin chains on it. The K63-linked poly-ubiquitin chains work as scaffolds for the assembling of protein kinase complexes and trigger their activation. Then the TAK1, which is one of the MAP3Ks, is recruited and activated by the poly-ubiquitinated TRAF6. After that, TAK1 activates downstream pathways, like JNK/p38 cascades[23]. But TAK1 is not the only mediator of TGFβ-induced JNK/p38.

MEKK1 and MLK3 have also been proposed to be able to activate JNK or p38 MAPK through MKK4 or MKK3/6, respectively[22].

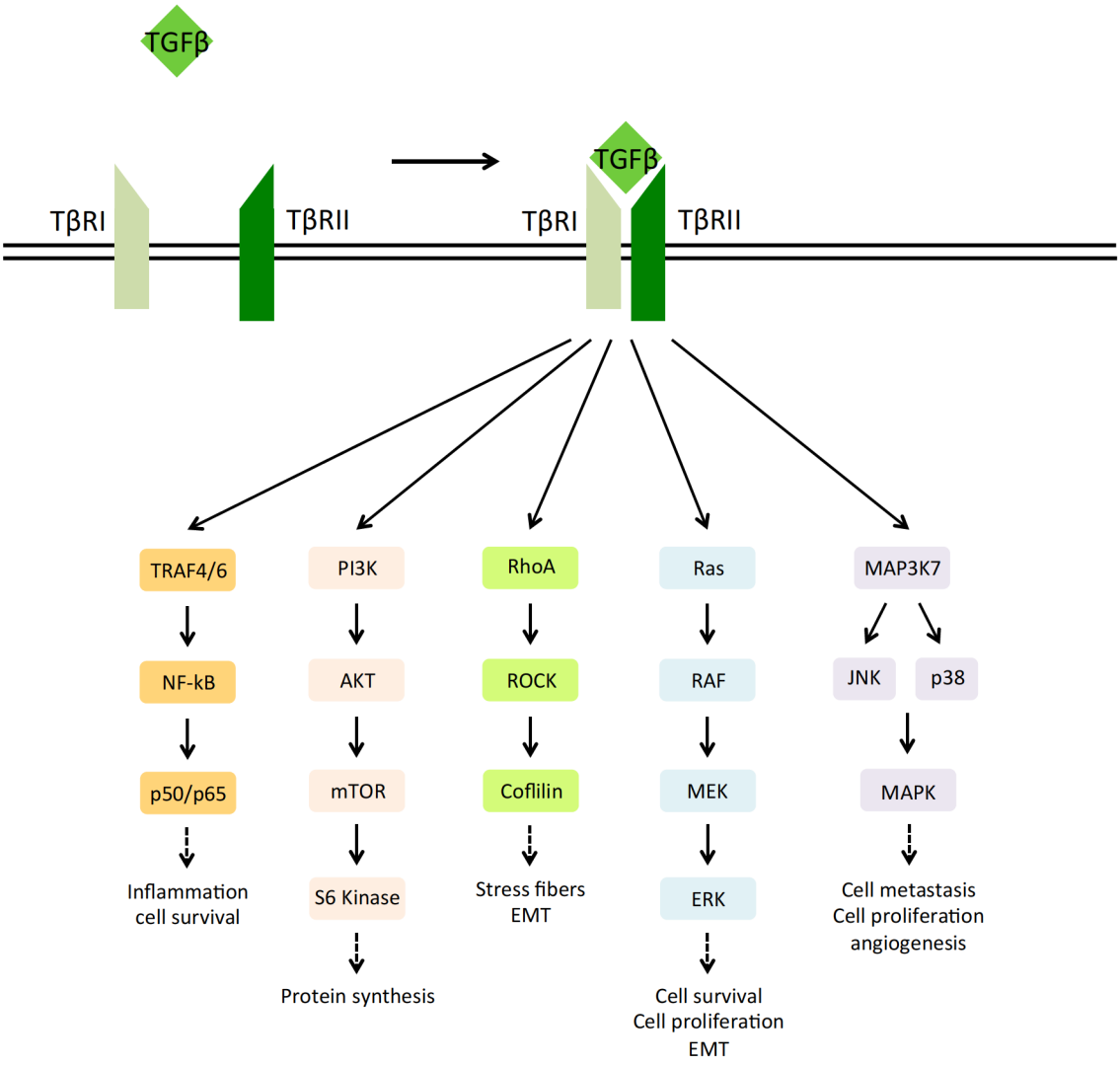


Figure 1.2 Diagram of non-canonical TGFβ signaling pathway. In the non-canoical TGFβ signaling pathway, the signal transmitted by several factors, including TRAF4/6, PI3K, RhoA, Ras and MAP3K7. The figure is adapted from Costanza, B (2017)[21].

Although tremendous progresses have been made to understand the TGFβ signaling networks in the past few years, it is still not clear how the cooperation between canonical and non-canonical signaling pathways leads to the determination of final cell fate.

1.2 Growth control by TGFβ

Among the fundamental functions of TGFβ in regulating cellular processes, growth control is considered in the clinical applications.

1.2.1 The proliferation control by TGFβ in epithelial and non-epithelial cells

Numerous studies indicate the crucial role of TGFβ in growth inhibition of epithelial cells, including cells from skin, colon, liver, mammary, pancreas and prostate[9]. Several developmental stages of mammary gland are influenced by TGFβ (Figure 1.3, upper left). During pregnancy and lactation, the branching ductal network of virgin mammary gland undergoes proliferation and differentiation[24-26]. During the involution, the mammary gland undergoes apoptosis and tissue remodeling to return to the virgin form[27]. These steps are regulated by TGFβ. The skin is another organ that is sensitive to TGFβ (Figure 1.3, bottom left). Interestingly, the effects of TGFβ on keratinocyte are contradictory, depending on the degree of cell differentiation and the level or the duration of TGFβ signaling. For example, in the differentiated suprabasal layers, TGFβ induces mitosis[28]. However, it causes growth inhibition in the basal layer[29,30].

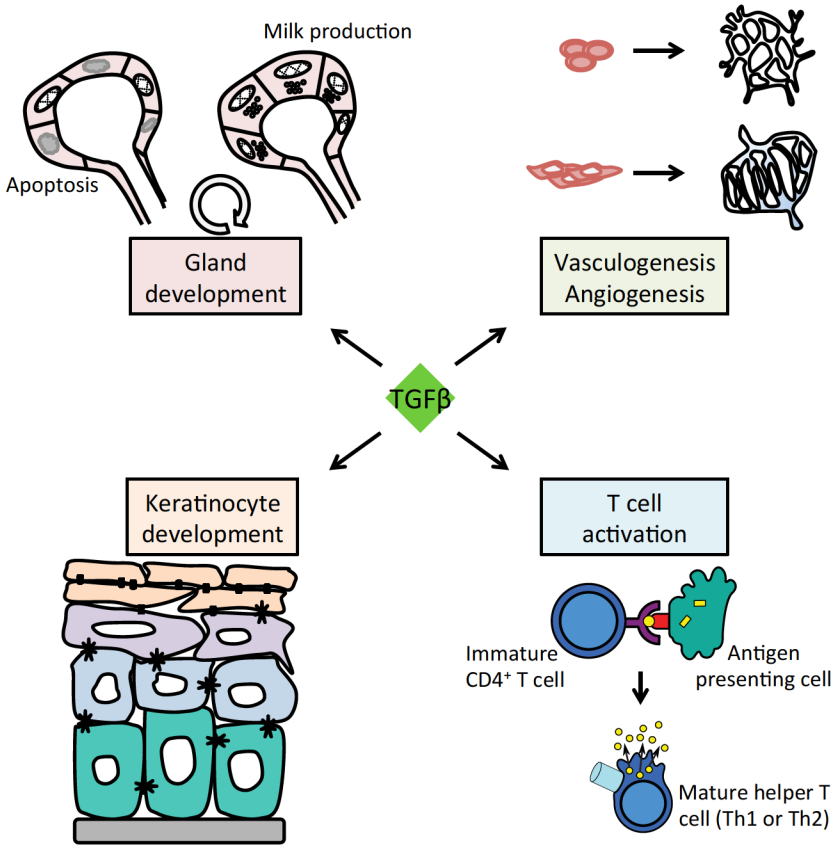


Figure 1.3 TGFβ controls the proliferation of epithelial and non-epithelial cells. Upper left: Several developmental stages of mammary gland are influenced by TGFβ during pregnancy, lactation and involution. Bottom left: The effects of TGFβ on keratinocyte are contradictory: in differentiated suprabasal layers, TGFβ induces mitosis. However, it causes growth inhibition in the basal layer. Upper right: In endothelial systems, the vasculogenesis and angiogenesis require the involvement of TGFβ signaling. Bottom right: In the immune system, TGFβ has several effects on the T lymphocytes, such as preventing the IL-2 production and T-cell activation.

TGFβ also controls the proliferation of many non-epithelial cells (Figure 1.3, right), such as immune, endothelial, and neuronal systems. Vasculogenesis and angiogenesis require the involvement of TGFβ signaling[31], though TGFβ inhibits growth and induces apoptosis of endothelial cells[32]. The opposite effects are conducted through the activation of two isoforms of TGFβ type I receptors, which are ALK1 and ALK5. ALK1 transduces signal via Smad1 and induces proliferation, whereas ALK5 activates Smad2/3 to inhibit cell growth[33]. These two pathway activations occur in response to different ligand concentrations, which provide basis for the differential responses. Immune system is another target of TGFβ-mediated proliferation control. There are several effects of TGFβ on T lymphocytes, such as preventing the interleukin-2 (IL-2) production[34], inhibiting T-cell activation, inducing peripheral T-cell apoptosis[35] and inhibiting the differentiation of T-helper1/2[36].

1.2.2 Mechanisms of TGFβ cytostatic effect in the cell cycle regulation

Most studies on the mechanisms of TGFβ cytostatic effect are conducted in epithelial cells, but some basic observations are also confirmed in other cell types. Cytostatic genes are activated by TGFβ at any time point in the G1, S, G2 phase during cell cycle, but they primarily target events in G1 phase[9]. Reversible G1 arrest induced by TGFβ has been studied in lung, skin and mammary epithelial cells.

The cell cycle is tightly controlled by two groups of proteins: Cyclins and Cyclin-dependent kinases (CDKs). Four Cyclin proteins, called CyclinA, CyclinB, CyclinD and CyclinE, fluctuate in predictable patterns throughout the cell cycle (Figure 1.4).

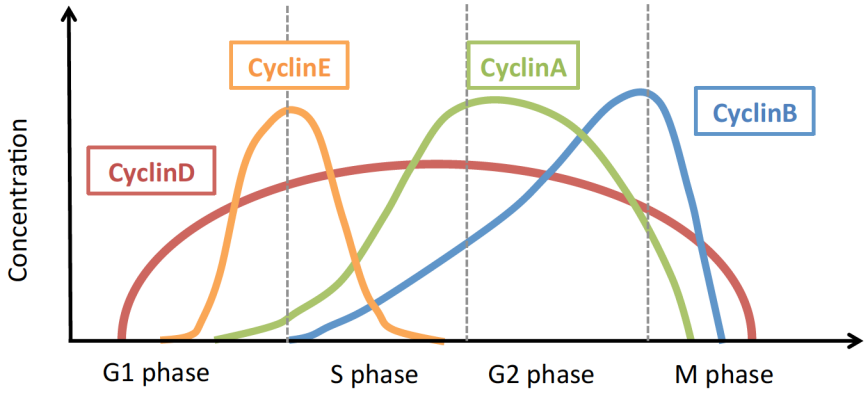


Figure 1.4 The fluctuation of Cyclin proteins throughout the cell cycle. The Cyclin protein accumulations correlate to the cell cycle checkpoints (G1/S, S/G2 and G2/M transition). The figure is adapted from Morgan D (2006)[37].

Cyclin proteins bound to CDKs and form complex. Then the complex is phosphorylated and activated, advancing the cell to the next cell cycle phase[38]. Throughout the cell cycle, the levels of CDKs are relatively stable, while the levels of Cyclin proteins fluctuate and determine the formation of Cyclin/CDK complexes. The specific concentrations of activated Cyclin/CDK complexes trigger the transitions to next phases[37].

The cytostatic programme of TGFβ includes the activation of CDK inhibitors (p15 and p21), and the repression of growth-promoting transcription factors (ID1, ID2, ID3 and c-Myc)[9]. By this programme, TGFβ induces cell cycle arrest at the G1 phase (Figure 1.5).

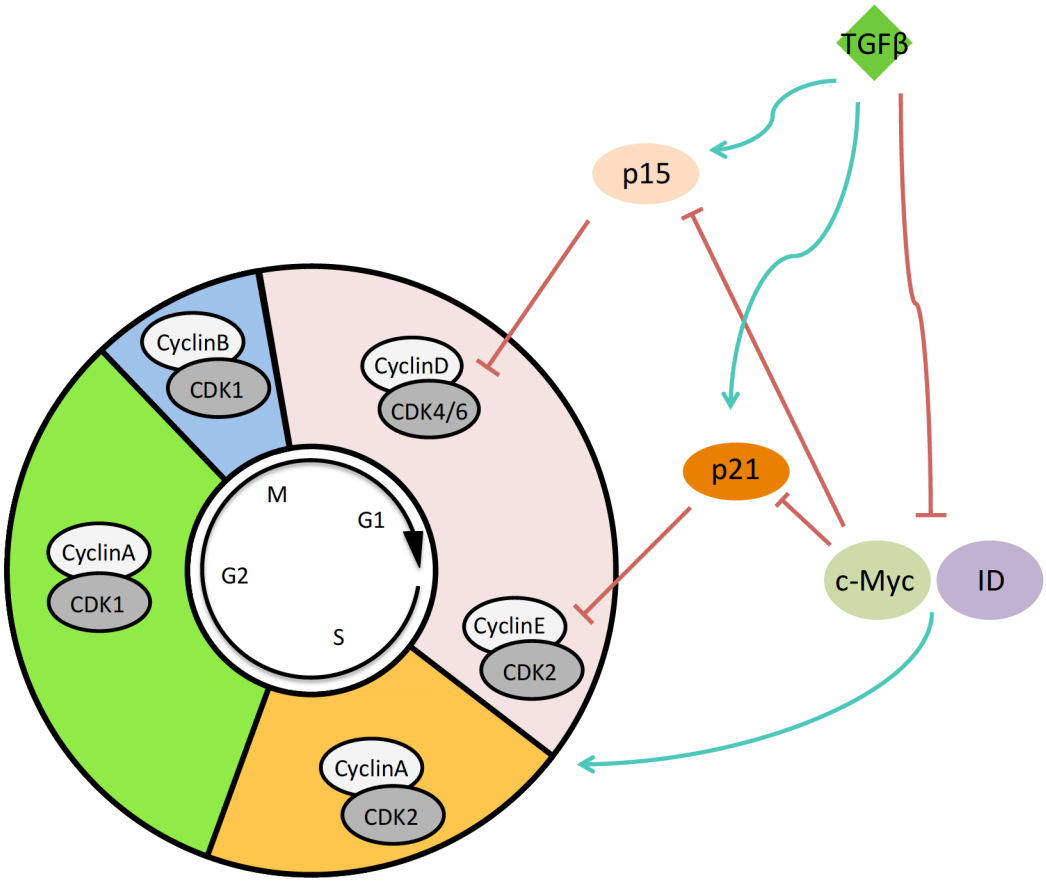


Figure 1.5 The TGFβ cytostatic programme. TGFβ controls the G1/S transition by up-regulating CDK inhibitors (p15 and p21) and down-regulating growth-promoting transcription factors (c-Myc and ID family).

The progression of G1 phase is driven by two complexes: CyclinD-CDK4/6 and CyclinE-CDK2[39]. p15 can directly interact with CDK4 and CDK6, inhibiting their Cyclin binding affinity by disrupting the catalytic sites[40]. p21 inactivates the

CyclinE-CDK2 complex by binding to the CyclinE subunit then blocking the CDK2 catalytic site, which results in the kinase inhibition[41-43].

In growing cells, c-Myc binds to the proximal region of the promoters of gene *CDKN2B* and *CDKN1A*, which encode p15 and p21, respectively[44]. Thus the transcription of these two genes is inhibited. TGFβ activates p15 and p21 by combination of removing transcriptional repression and triggering transcriptional activation. On one hand, TGFβ down-regulates c-Myc and releases it from the promoters of *CDKN2B* and *CDKN1A*[45]. On the other hand, the TGFβ activated Smads complexes bind to these promoters, cooperate with DNA-binding co-factors and activate the transcription.

ID proteins promote cell proliferation, which are down regulated by TGFβ. TGFβ can induce the transcription of activating transcription factor-3 (ATF3), which interacts with Smad protein complex to repress the transcription of *ID1* [18]. For *ID2*, c-Myc can bind to the E-box motifs of its promoter to enhance its transcription. Thus TGFβ-induced c-Myc inhibition leads to the down-regulation of ID2[46].

Taken together, the ability of TGFβ to inhibit Cyclin-CDK activity appears to be its strategy to block G1/S transition and control cell proliferation.

1.3 Single cell behaviors are masked by measurements on population

In the past decades, there is an increasing interest in the heterogeneous cellular behaviors. The heterogeneous responses may be caused by inherent stochasticity in the biochemical processes[47], the variation in the numbers of proteins in different individual cells and spatially polarized signaling networks[48]. Thus individual cells response differently to uniform physiological stimuli[49].

Although the population averaged assays are powerful tools in biological researches, their limitation of losing information of cell heterogeneity is obvious[48]. Several patterns of single cell behavior can generate the same population behavior when averaged (Figure 1.6).

It was difficult to measure the level and activity of network components in single cells decades ago. Fortunately, the rapid development of technologies allows more precise manipulation and detailed observations at the single cell resolution. For example, the flow cytometry makes it possible to access the distribution of single cell behaviors. Combining flow cytometry with fluorescence *in situ* hybridization, it is possible to detect mRNA and protein in

single cells[50-53]. With microscopy, we can monitor the morphological characteristics of individual cells. Single cell sequencing provides a high resolution of measuring the gene expression profiles of heterogeneous populations[54]. Live cell imaging allows us to analyze single proliferating cells over time[55].

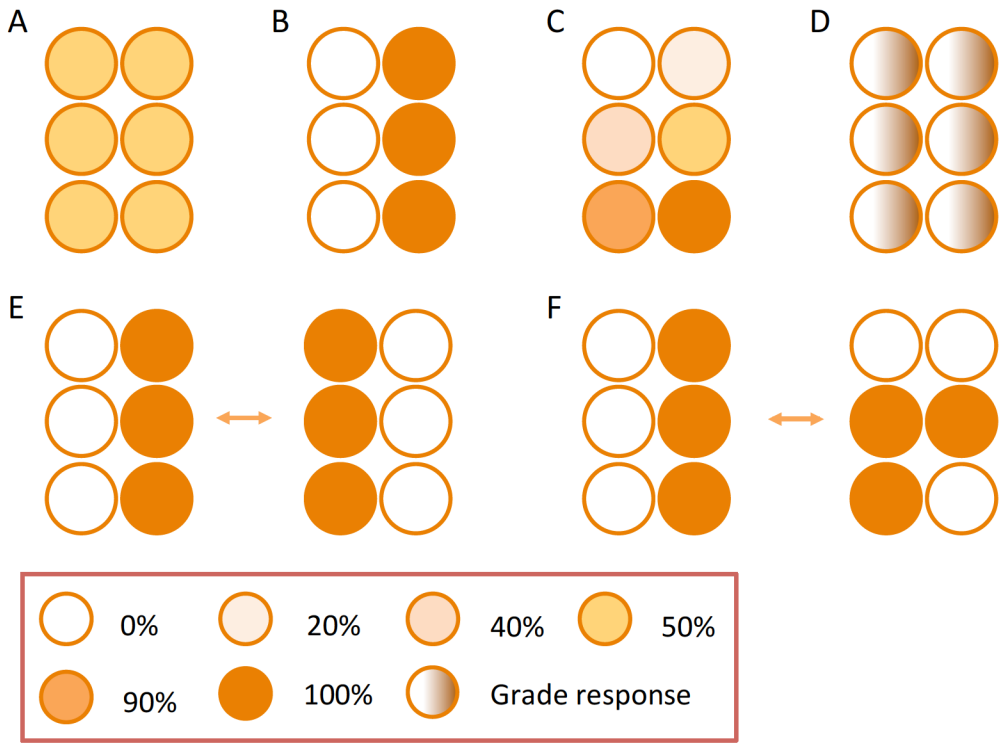


Figure 1.6 Several patterns of single cell behavior generate the same population behavior on average. Cells are shown as circles and their signaling activities are indicated as orange color. The average signaling in each panel is 50%. (A) All the cells signal at 50%. (B) Half of the cells signal at 0% and half of the cells signal at 100%. (C) Signals of the cells are distributed from 0% to 100%. (D) Each cell displays spatial heterogeneous signal. (E) Cells are either signal at 0% or 100%, and they change into the other state at specific time, without influence the population average. (F) Cells are either signal at 0% or 100%, and they change to the other state stochastically. The figure is adapted from Cheong *et al* (2010)[48].

The growth inhibitory function of TGFβ has been extensively studied. However, the knowledge that TGFβ induce cell arrest in G1 phase is based on the observation of averaged population behaviors at specific time points[56-59]. The exact response in individual cells is not clear. For instance, it is not known how long a cell arrests; whether all cells or just partial of them are affected; and how the response differs upon different ligand doses or stimulation time at the single cell level.

In addition, to investigate phase-dependent behaviors at the population level, physical or chemical synchronization were conducted[60,61]. In general, when

quiescent cells start to recover from the un-physiological state, the cell proliferation rate differs from that under normal growth conditions, as additional metabolic reactions are required to return to the cell cycle[62]. Besides, methods based on DNA replication inhibition to achieve cell synchronization cause DNA damage response[63]. Thus synchronization introduces artifacts and may mask the natural cellular responses. Moreover, the achieved synchrony is usually lost in one cycle, which makes it difficult to track inherited effect over generations.

Hence, quantitative analysis on asynchronous single cells is crucial to understand how the individual cells decode and integrate TGFβ signal into proliferation decisions.

1.4 Quantitatively measuring the signaling dynamics in single cells

Signaling networks transmit information of the extracellular environment to the downstream effectors in the cell[64]. Then the cell integrates the information and reacts by changing physiological state. Due to the inherent stochasticity of biochemical reactions in signaling transduction (intrinsic noise) and the variability of cellular states upon stimulation (extrinsic noise), the capacity of cellular signaling pathways is limited[65]. There is an emerging trend suggesting that information is transmitted in cells by the dynamics of signaling molecules. It was demonstrated by theoretical analysis that the signaling dynamics play an important role in eliminating extrinsic noise and precise information transmission[66]. To understand cellular information encoding and heterogeneous responses, it is vital to measure the molecular processes quantitatively and dynamically[67].

1.4.1 Quantifying the dynamics of cellular molecules and temporal signaling

There are several modes of signaling in response to stimulus in cells: the change in abundance (e.g. protein expression is induced by stimulus), identity (e.g. phosphorylation, ubiquitinoylation or binding to a cofactor), and location (e.g. translocation from the cytoplasm into the nucleus)[68] (Figure 1.7A, B and C). The dynamics of cellular molecules can be defined as the shape of curve that describes the change of a molecule in these aspects over time (Figure 1.7D).

The static characters (changes in abundance, identity or location at specific time) show the information of the molecular at only one single point in time. The temporal signal displays the molecular dynamics during a period of time, the information of which is richer and more complex[69].

For the temporal signal, the information can be encoded as features of the signal curve, including delay, amplitude, frequency, duration and cumulative level (Figure 1.8). The cell decodes and integrates extracellular signals, then responses to the changes in the environment.

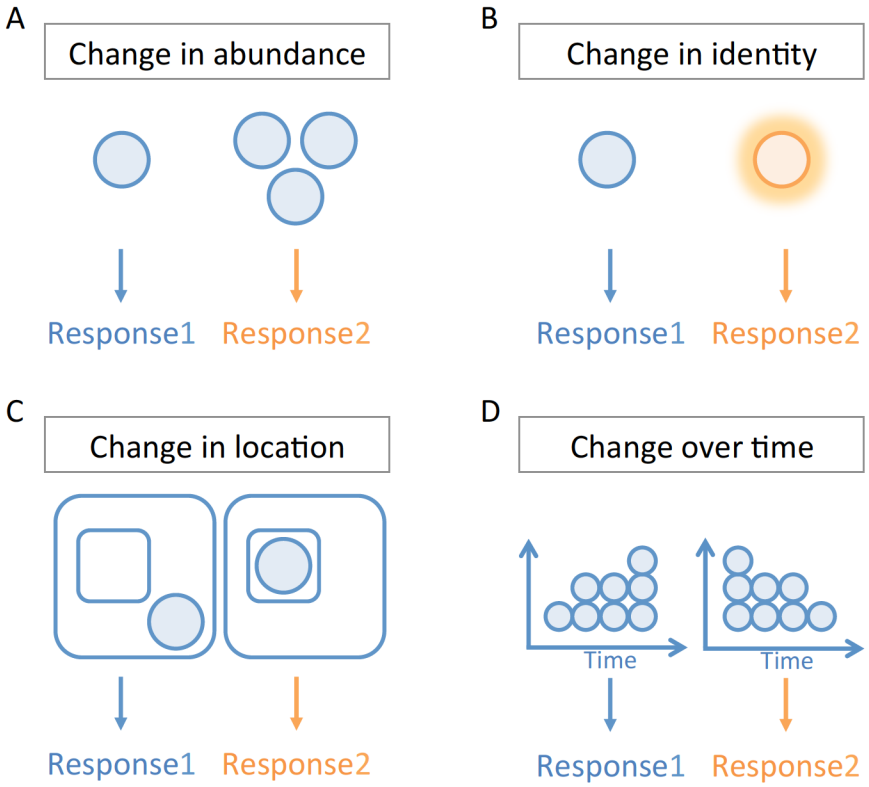


Figure 1.7 Quantifying the cellular molecular dynamics. There are several modes of signaling in response to stimulus. Different signals determine different responses. The signal can be distinguished by quantities: (A) molecular abundance, (B) identity, and (C) location. (D) The dynamics can be defined as the curve describing the temporal pattern of signal. The figure is adapted from Jeremy E. Purvis *et al* (2013)[70].

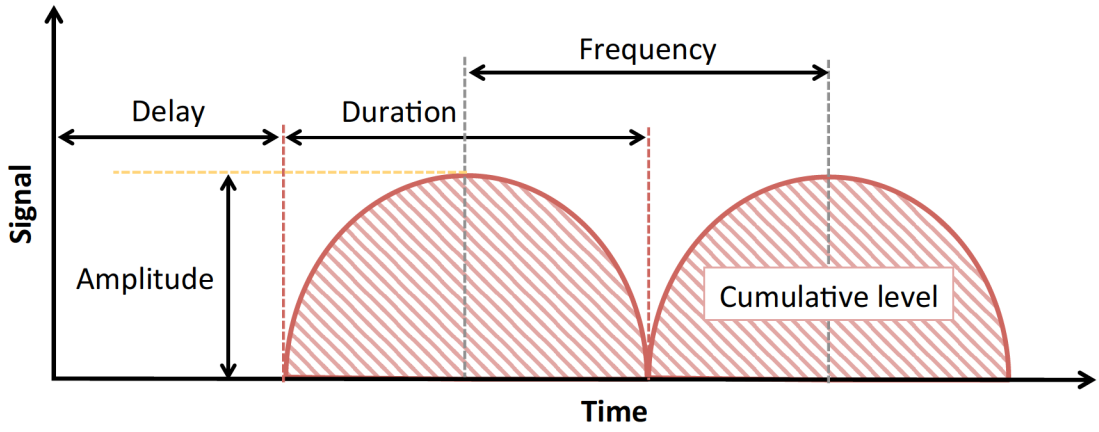


Figure 1.8 The measureable features of temporal signal. The information can be encoded as the features of temporal signal curve, such as delay, amplitude, duration, frequency or cumulative level. The figure is adapted from Jeremy E. Purvis *et al* (2013)[70].

1.4.2 Techniques that are developed for measuring cell signaling dynamics

There are several experimental tools to measure cellular signal dynamics[68]. For examples, the patch-clamp method can be used to measure currents passing one ion channel[71,72]; RNA fluorescent *in situ* hybridization (RNA FISH) provides evidences for dynamic and stochastic transcription[73]; Stimulated Raman Scattering (SRS) and Coherent Anti-Stokes Raman Scattering (CARS) are used in specific molecular detection in cells or tissues[74,75].

The most powerful tools are the techniques for tracking intracellular proteins by taking advantage of fluorescent proteins. The rapid development of these techniques results in considerable progress in quantitatively analysis on the spatiotemporal signaling dynamics in individual cells[76,77]. Here, we list several techniques for measuring protein translocation, interactions and modifications (Table 1-2).

Table 1-2 Techniques for measuring intracellular proteins dynamics

Method	Examples of application	Reference
Fluorescence recovery after photobleaching (FRAP) and fluorescence loss in photobleaching (FLIP)	Measuring the rate of protein movement	78
Photoactivation and photoconversion	Measuring protein turnover rate	79
Fluorescence correlation spectroscopy (FCS)	Measuring the absolute number of proteins in the confocal volume	80
Image correlation spectroscopy (ICS)	Time-lapse imaging	81
Fluorescence cross-correlation spectroscopy (FCCS)	Detecting the binding relationships between molecules	
Image cross-correlation spectroscopy (ICCS)	Detecting protein-protein interaction	81
Fluorescence resonance energy transfer (FRET)	Detecting if two proteins are in close proximity (1-10nm)	82
Fluorescence lifetime imaging microscopy (FLIM)	Measuring the rate of donor emission decay	83
Protein-fragment complementation assay (PCA)	Detecting protein-protein interaction with low background	84-86

1.4.3 Monitoring Smad2 translocation to quantify TGFβ signaling dynamics

With the development of fluorescent time-lapse microscopy, several signaling events have been studied thoroughly, such as p53 and NF-κB signaling (Figure 1.9), which highlight the correlation between signaling dynamical patterns and specific downstream responses[69].

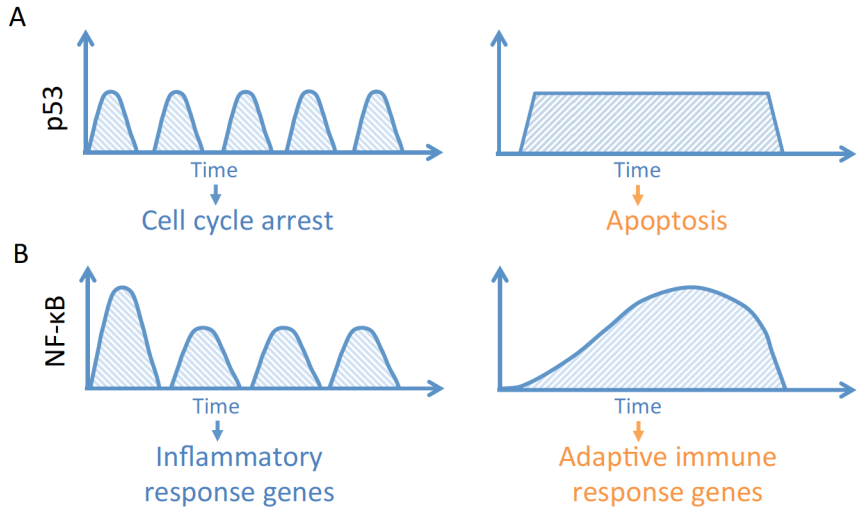


Figure 1.9 Different signaling dynamics lead to different downstream responses. (A) p53 pulses lead to cell cycle arrest and prolonged p53 signaling leads to apoptosis. (B) Transient NF-κB nuclear accumulation triggers the expression of inflammatory genes and sustained nuclear NF-κB level triggers the expression of specific genes for adaptive immune response. The figure is adapted from Jeremy E. Purvis *et al* (2013)[70].

In a fluorescent system to study the p53 signaling, the fusion p53-Venus is expressed in the cell and the activation of p53 is indicated by the Venus intensity[87,88]. In the NF-κB signaling pathway, activated NF-κB translocates from cytoplasm into the nucleus. Thus in the fluorescent system that expresses fusion NF-κB (p65/RelA), the oscillations of NF-κB nuclear-cytoplasmic shuttling can be quantitatively measured to indicate signaling activation[89].

In this thesis work, we construct a fluorescent system to study TGFβ signaling in single cells. In the canonical TGFβ signaling pathway, several key steps are involved: activation of receptors (TβRI or TβRII), cytoplasmic-nuclear shuttling of R-Smads (Smad2 and Smad3) and transcription of TGFβ downstream target genes (Figure 1.1).

Due to the cell-cell contact and adhesion, it is difficult to quantitatively measure the abundance of receptors on the plasma membrane in each individual

cells. TGFβ can activate more than 300 target genes[9]. One of them cannot represent the primary activities of TGFβ signaling.

The cytoplasmic-nuclear shuttling of Smad2/3 is the best indicator for quantitatively measuring TGFβ signaling dynamics. Similar to NF-κB activation, the cytoplasmic-nuclear shuttling (N-C shuttling) of R-Smads can be easily monitored and quantitatively measured[90,91]. The N-C shuttling patterns of Smad2 and Smad3 are similar. Thus we decide to employ fluorescent Smad2 fusion to study TGFβ signaling. The Smad2 signaling dynamics is TGFβ dose-dependent and highly heterogeneous[90]. We don't have a clear understanding on the correlation between Smad2 dynamics and downstream response in individual cells (Figure 1.10).

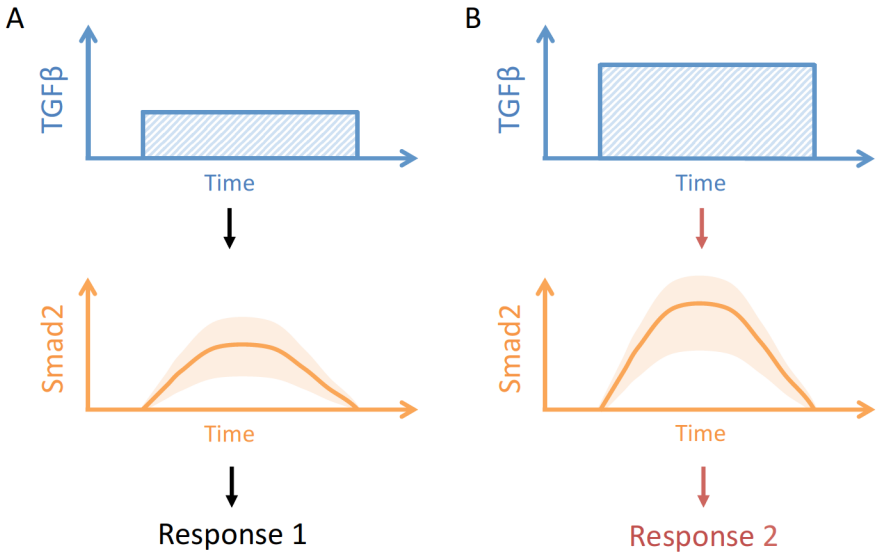


Figure 1.10 Examples of the Smad2 dynamics in response to TGFβ. Different nuclear accumulations of Smad2 are triggered by low dose (A) or high dose (B) of TGFβ.

One of the crucial downstream responses to TGFβ is cell cycle arrest. The exact outcomes of cell cycle arrest in individual cells are not clear. In order to study the signal transmission from TGFβ to Smads and finally to the cell division decisions, we need another fluorescent reporter for cell cycle progression.

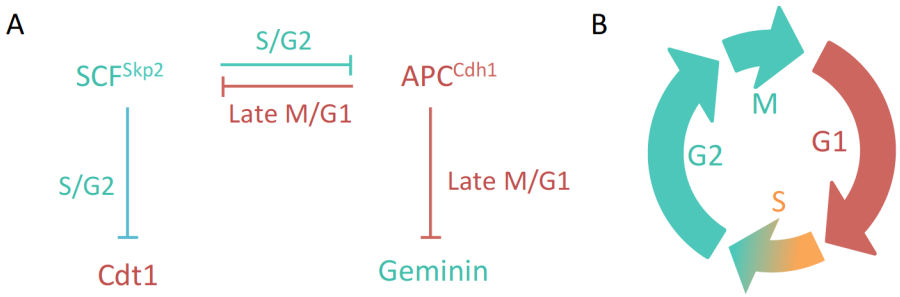
1.5 Fluorescent cell cycle indicator (FUCCI)

There are several available fluorescent cell cycle indicators for live cell. For instance, mVenus-p27K$^-$ is used as an indicator for G0/G1 transition[92]; the S phase

could be recognized by the distinct replication structures of mCherry-PCNA[93-96]; H2B-GFP makes the chromosome condensation in mitotic progression visible[97].

One of the most widely used cell cycle indicators is FUCCI (Fluorescence Ubiquitin Cell Cycle Indicator), which is accurate and versatile[98,99]. Since the introduction of FUCCI in 2008, this technology revolutionizes the analysis on cell proliferation and leads to several important discoveries[100,101].

The molecular mechanism of the first generation of FUCCI involves two components in the DNA replication control system: Cdt1 and Geminin. These two proteins have opposing effects on the DNA replication and their abundances oscillate oppositely through the cell cycle. Cdt1 is the licensing factor of replication and it peaks in G1 phase. Geminin is the inhibitor of Cdt1 to ensure the replication occurs only once in one cell cycle. The protein level of Geminin oscillates and peaks during S/G2/M phase[102]. The oscillations of Cdt1 and Geminin are controlled by SCF^{Skp2} and APC^{Cdh1}, respectively (Figure 1.11 A and B).



C

	AmCyan	mAG	Venus	mKO2	mCherry
$hCdt1_{(30/120)}$	-	-	-	+	+
$hGem_{(1/110)}$	+	+	+	+	+
$hGem_{(1/60)}$	?	+	+	+	+

Figure 1.11 Characterization of FUCCI. (A) The oscillations of Cdt1 and Geminin are controlled by SCF^{Skp2} and APC^{Cdh1}. (B) The fluorescent signal of original FUCCI (mKO-$Cdt1_{(30/120)}$ (red) and mAG-$hGem_{(1/110)}$) (green) in cell cycle. (C) Fluorescent proteins in functional FUCCI. The figure is adapted from Asako Sakaue-Sawano *et al* (2008)[103].

Only the truncated Cdt1 and Geminin proteins are used in FUCCI: $Cdt1_{(30/120)}$ and $hGem_{(1/110)}$ are constructed to indicate G1 and S/G2/M phases, respectively. They are sufficient for indicating cell cycle phase, and do not interfere cell cycle progression. However, both of them locate in nucleus, which loses the information of cell morphology in cell proliferation and differentiation. Thus a shorter Geminin fragment ($hGem_{(1/60)}$) is constructed to overcome the shortage, which is evenly distributed throughout the nuclei and cytoplasm.

The original FUCCI is the system of mKO-$Cdt1_{(30/120)}$ and mAG-$hGem_{(1/110)}$. With the development of this tool, FUCCI with various fluorescent proteins are available (Figure 1.11 C), such as mCyan, Venus or mCherry. The increased selections of fluorescent proteins expand the studies that correlate cell cycle with other cellular processes, for example, signaling events.

In this project, we combine FUCCI (mCherry-$hGem_{(1/110)}$) with EYFP-Smad2 to investigate the cell cycle regulation by TGFβ/Smad signaling.

1.6 Mathematical modeling is a powerful tool in studying cell cycle progression

In modern biology, computational or mathematical modeling plays an important role in biological research. The systematical perspective to investigate biological behaviors drives constant innovation in life science. On one hand, the application of dynamical systems theory to the biological processes facilitates the research by providing spatiotemporal insights. On the other hand, computational simulation inspires the efficient experimental designs by revealing emergent properties.

In this project, we would like to employ mathematical modeling to simulate cell cycle progression, guiding the experimental design and analysis.

In systems biology, one of the central elements is the interplay between mathematical modeling and experiments. In the field of modeling cell cycle, there are two main camps of mathematical models, one is at mechanistic level, the other is at the descriptive level[104]. Mechanistic models represent protein interaction networks, e.g. Cyclins and CDKs interactions, to explain the underlying molecular mechanisms of cell proliferation[105,106]. The descriptive models focus on representing the characteristics of observable cause-effect phenomena in cell cycle progression (e.g. size, cell cycle phase and cell number), to explain the overall system behaviors[107-110].

1.6.1 Mechanistic network models of cell cycle

Many mechanistic models, which include molecular mechanistic details underlying cell cycle progression, have been proposed and developed. During the past decades, significant progress has been made in the identification of cell cycle related genes and their regulatory networks[111]. A large number of molecules have been demonstrated to be involved in the regulation of cell cycle progression[112,113]. The central elements for controlling cell cycle progression are complexes of Cyclin proteins and CDKs. Checkpoint mechanisms govern the irreversible phase transitions, which ensures normal cell growth and mitosis[114,115].

The accumulation of knowledge on molecules involved in cell cycle stimulates the development of mathematical modeling research on cell cycle. One of the earliest works is the model investigating the regulatory mechanism that controls the M phase entry[116]. This model develops a set of ordinary differential equations (ODEs) based on experimental data of *Xenopus* oocyte extracts. Analysis on this model reveals that positive feedback loops play an essential role in the switch of mitosis entry. This model also predicts the bistability of the cell cycle control system. Late on, the existence of bistable switch is verified by experiments[117,118]. In the mammalian system, various models have been proposed to explain the phase transitions in the cell cycle[119,120], such as G1/S transition[121-124], and G2/M transition[125] (Figure 1.12).

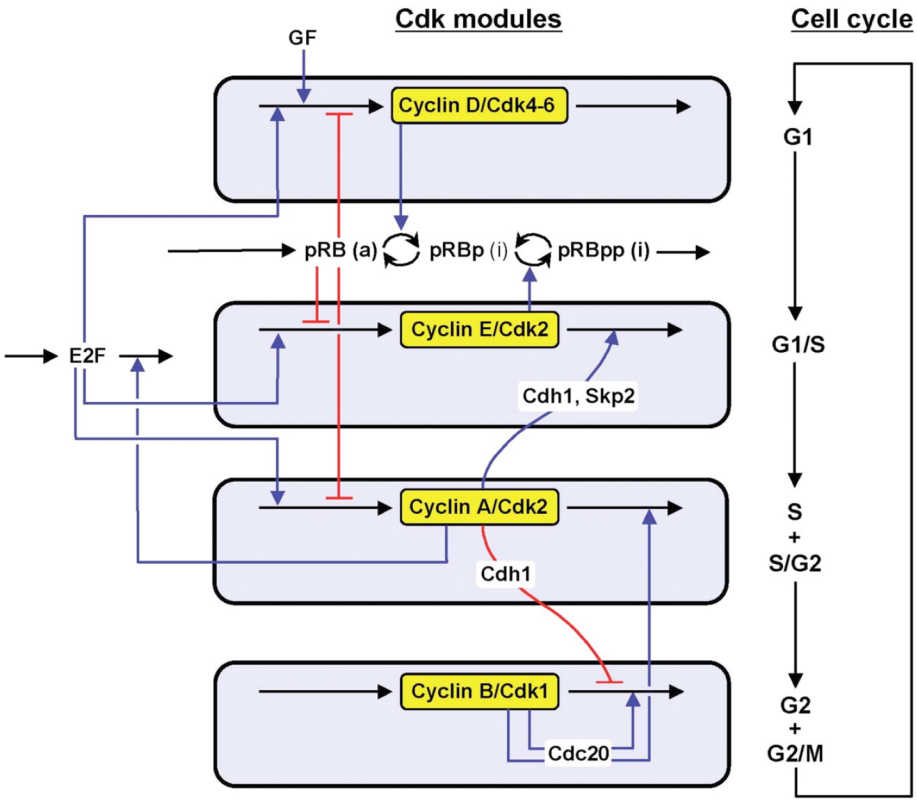


Figure 1.12 A model of the mammalian cell cycle. This model incorporates the G1/S, S/G2 and G2/M phase transitions. Four CDK modules, CyclinD/CDK4-6, CyclinE/CDK2, CyclinA/CDK2 and CyclinB/CDK1, control the cell cycle progression. The effects of GF and pRB/E2F are also involved in the regulation. The figure is taken from Claude Gerard, *et al* (2009)[119].

1.6.2 Descriptive network models of cell cycle

In contrast to the mechanistic models, the descriptive models, which focus on the overall cellular behaviors instead of mechanistic details, have been developed and applied to study the variability in cell cycle progression. There are many descriptive models accounting for cell cycle phase transitions, such as branching processes, transition probability models (TP models) and growth control models[126]. There is a long history of applying branching processes to biological systems. One derivative of the application is to analyze the cell lineage data: for the cells in one generation, each of them produces two newborn cells in the next generation[127,128]. The TP model is another example of phenomenological model, which was developed around 40 years ago[129-132]. In the TP model, the variability of cell division time originates from random transitions through different cell cycle phases. The advantage of TP models is that they have only a few parameters. Although the molecular mechanism is not provided in TP models, these models have been successfully applied to a number of experimental data, providing useful information on cell cycle progression (Figure 1.13).

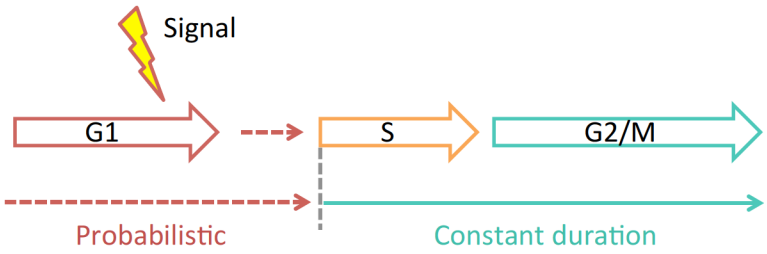


Figure 1.13 Schematic diagram shows the original transition probability model. Once the cell receives the signal, e.g. growth factor stimulation, the cell progresses the G1/S transition stochastically with a value of probability. After a constant duration, the cell divides. The figure is adapted from Takashi Saitou *et al* (2015).

Both mechanistic and descriptive models are useful in the studies of cell cycle. For specific issues, we should choose the proper mathematical model and refine it to adapt to specific questions.

1.7 Motivation and outline of this thesis

One of the most fundamental functions of TGFβ is growth inhibitory effect. TGFβ induces cell-cycle arrest in G1 phase. However, this knowledge is based on the observation of averaged population behaviors at specific time points[56-59]. The exact response in individual cells is not clear.

In this project, combining model prediction, time-lapse microscopy and single cell lineage analysis, we would like to study G1 arrest induced by TGFβ in asynchronously proliferating single cells and answer the following main questions:

1) Several different patterns of single cell behavior can generate the same population behavior. Cell could respond to TGFβ either in a **graded** mode (a graded slow-down of progression in G1 phase is triggered by increased doses of TGFβ and the arrested time is adjusted according to the concentration), or in an **all-or-none** mode (partial cells are arrested in G1 phase and the affected fraction is proportional to the ligand dose). Which response mode is correct in single cells?
2) Asynchronously proliferating cells are in different cell cycle phases upon TGFβ stimulation. Is the response to TGFβ cell cycle phase dependent?
3) Asynchronously proliferating cells are in different stages within each cell cycle phase upon TGFβ stimulation. Is the response to TGFβ dependent on the timing of treatment within the phase?
4) The Smad signaling plays a central role in the canonical TGFβ signaling pathway. Does the Smad2 signaling dynamics determine the G1 arrest duration? How does the cell decode and integrate the information of TGFβ stimulation, then make the final decision on cell cycle phase progression?
5) Single cell data shows important information. But the single cell data is more difficult to achieve than population data. Is it possible to bridge the gap between the population and single cell scales by mathematical modeling? Can we use modeling to facilitate inferring single cell information based on the easily accessible population data?

This work characterizes the proliferation decisions in response to TGFβ and provides tools to investigate single cell behavior, benefiting future studies on cytostatic responses, as well as the development of therapeutical strategies for cell proliferation control.

Chapter 2

A live-cell system to monitor the TGFβ signaling and cell-cycle progression

In this chapter, we will introduce the details of the fluorescent system used in this project. The initial questions and hypotheses are proposed.

To study TGFβ signaling and cell cycle simultaneously, we constructed a cell line expressing EYFP-Smad2 and FUCCI reporter to monitor TGFβ/Smad signaling and cell cycle progression in individual cells.

With time-lapse microscopy, we recorded the signaling dynamics and cell cycle progression in growing cells. The fractions of cells in different cell cycle phases (G1 or S/G2/M phase) over time were quantitatively and efficiently measured. When the cells were treated with TGFβ, the fraction of cells in G1 phase increased. To explain this observation, we proposed two different hypotheses of G1 arrest response modes in individual cells: (1) Graded response; (2) All-or-none response.

2.1 Construction of a fluorescent reporter system for monitoring TGFβ signaling and cell-cycle phase transitions

In order to simultaneously monitor TGFβ signaling and cell cycle progression, we established a HaCaT cell line stably expressing EYFP-Smad2 fusion protein, a previously described fluorescent cell cycle indicator, FUCCI and a nuclear marker, H2B-ECFP. Smad2 nuclear translocation is a key step in the canonical TGFβ signaling, and thus was used to report the TGFβ signaling activity in this work.

In the original FUCCI system, mKO2-hCdt1(30/120) and mAG-hGem(1/110) were co-expressed to indicate G1 and the S/G2/M phases, respectively. Either of them is sufficient to distinguish G1 from S/G2/M phase, but co-expression of both facilitates tracking of proliferating cells. Nevertheless, when co-expressing both indicators in our cell line, the fluorescence of mAG-hGem(1/110) disappears rapidly in the late M phase but the fluorescence of mKO2-hCdt1(30/120) becomes detectable after a while, during this time the cells are totally dark. Though this dark gap is small, it still causes problems in cell tracking over generations. Thus we decided to use H2B-ECFP, a nuclear marker, for cell segmentation and tracking.

Taken compatibility of fluorescence spectra into account, we finally generated a stable HaCaT cell line co-expressing EYFP-Smad2, mCherry-hGeminin(1/110) and H2B-ECFP (Figure 2.1).

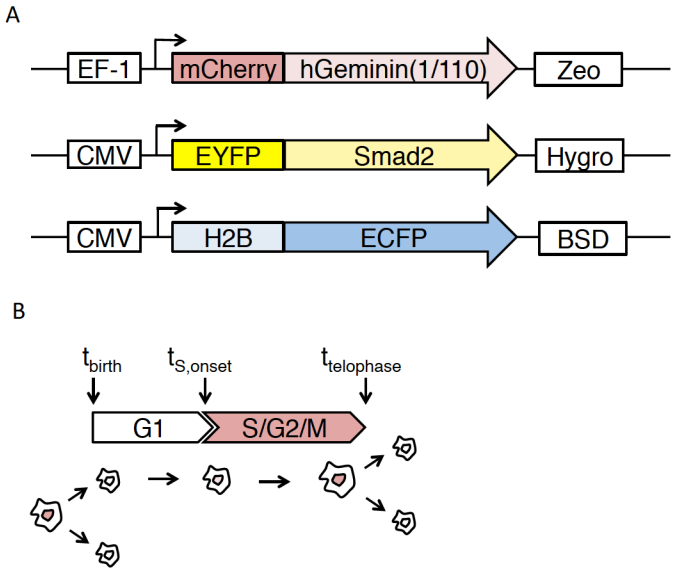


Figure 2.1 Schematic showing the design of fluorescent reporter system. (A) Construction of the fluorescent reporter system to monitor the cell cycle and Smad2 dynamics in individual cells. The FUCCI reporter, mCherry-hGeminin(1/110). The Smad2 was fused to Enhanced Yellow Fluorescent Protein (EYFP). The Histone 2B (H2B) was fused to Enhanced Cyan Fluorescent Protein (ECFP). (B) Cell cycle diagram showing the two phases (G1 and S/G2/M phases) indicated by different colors (no color and red color, respectively) in the system.

2.2 Validation of the fluorescent reporter system

2.2.1 The introduced fluorescent reporter system does not disturb cell cycle

To confirm that the reporters did not disturb cell cycle, we analyzed the DNA content of this reporter cell line by flow cytometer. The DNA content distribution of the reporter cell line is similar as the parental HaCaT cells (Figure 2.2), which confirms that the reporters do not interfere with normal cell cycle process.

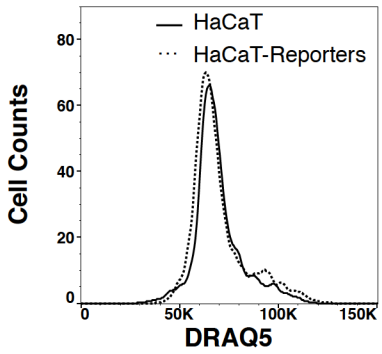


Figure 2.2 Analysis of DNA content in reporter cell line and the parental cell line. The DNA were stained by DRAQ5 and then analyzed by flow cytometer. HaCaT : Parental cell line; HaCaT-Reporters: The reporter cell line. Experiments were repeated twice independently with similar results.

2.2.2 The FUCCI reporter faithfully indicates the cell cycle phase

Next, we verified the functionality of the cell cycle indicator (mCherry-hGeminin(1/110)), i.e. whether only the cells with red nuclear signal are in S/G2/M phase (Figure 2.1B). After DRAQ5 staining, the DNA content and mCherry signal of this reporter cell line were measured by flow cytometer. We divided these cells into mCherry(+) and mCherry(-) clusters, and analyzed their respective DNA content by DRAQ5 staining signal. The mCherry(+) cells, which were supposed to be in S/G2/M phase, contained fully- and partially-replicated DNA (Figure 2.3).

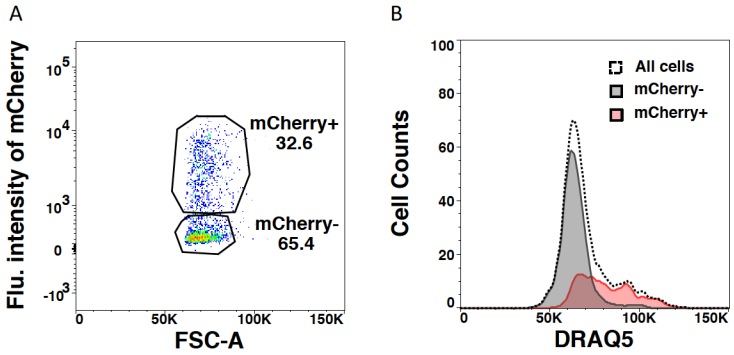


Figure 2.3 mCherry-hGeminin(1/110) functions effectively in the reporter cell line. The reporter cells showing red [mCherry(+)] and no fluorescence [mCherry(-)] were collected separately. Their DNA contents were stained with DRAQ5 and measured by flow cytometer. Experiments were repeated twice independently with similar results.

We also confirmed that the spatiotemporal properties of mCherry-hGeminin(1/110) under the microscope. When recording cell cycle progression by live cell imaging, we observed that the mCherry fluorescence disappeared after cell division and began to increase after around 8 hours (Figure 2.4A middle panel and C), which matches the typical G1 duration of this cell line. In addition, the red signal located in nuclei only. Thus, we could use the mCherry-hGeminin signal in the cell to indicate its S/G2/M phase.

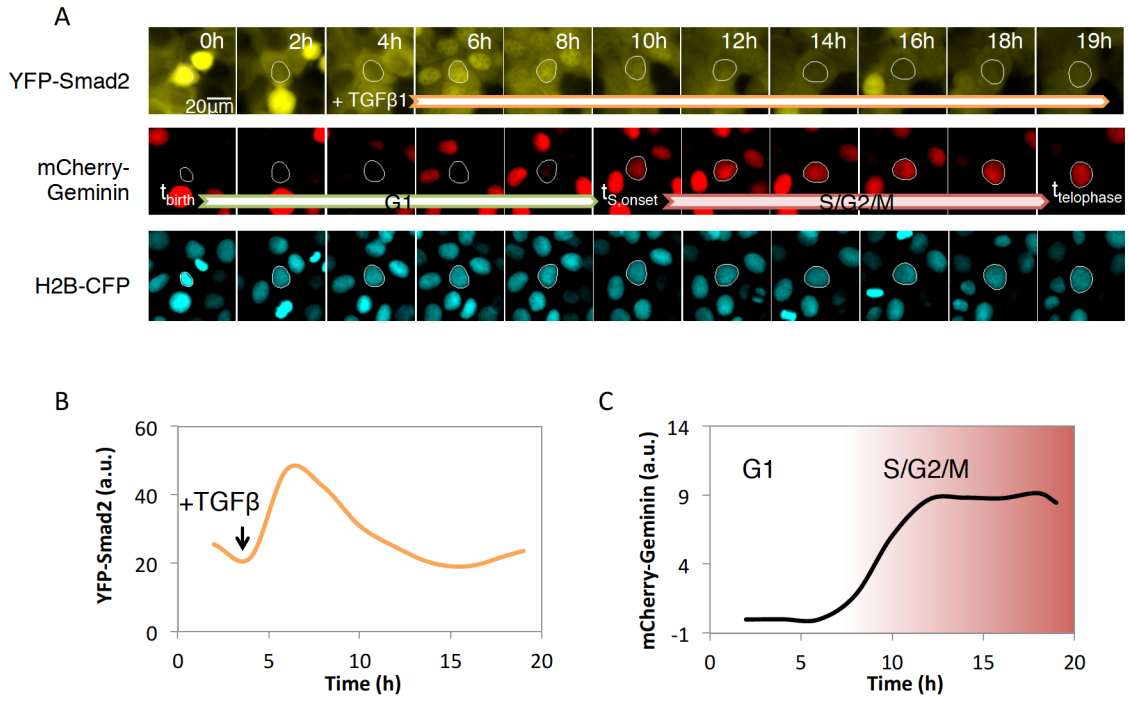


Figure 2.4 Measuring the dynamics of Smad2 and cell cycle progression in single cells. (A) The Representative live cell images of the HaCaT cells expressing EYFP-Smad2, mCherry-hGeminin(1/110) and H2B-ECFP. TGFβ is added at the 4h. The nuclear segmentation was indicated as white circle. The cell was tracked from its birth to its division. (B) Quantitative analysis of nuclear EYFP-Smad2 upon 100pM TGFβ1 stimulation of the cell in (A). (C) Quantitative analysis of nuclear mCherry-hGeminin(1/110) of the cell in (A). The onset of S phase was defined as the initial accumulation of red signal.

2.2.3 EYFP-Smad2 reflects endogenous Smad proteins dynamics

We used the dynamics of EYFP-Smad2 under the microscope to indicate the activation of TGFβ signaling. When monitoring the EYFP-Smad2 signal under the microscope, it was predominantly in the cytoplasm and a strong nuclear accumulation was observed upon TGFβ treatment (Figure 2.4A upper panel and B).

To validate if the EYFP-Smad2 could represent the endogenous Smad2 in TGFβ signaling transmission, we analyzed the Smad proteins dynamics in response to

100pM TGFβ1 stimulation. We demonstrated that the amount of EYFP-Smad2 is at a comparable level of the endogenous Smad2, and they had similar phosphorylation and translocation dynamic profiles in response to TGFβ stimulation (Figure 2.5).

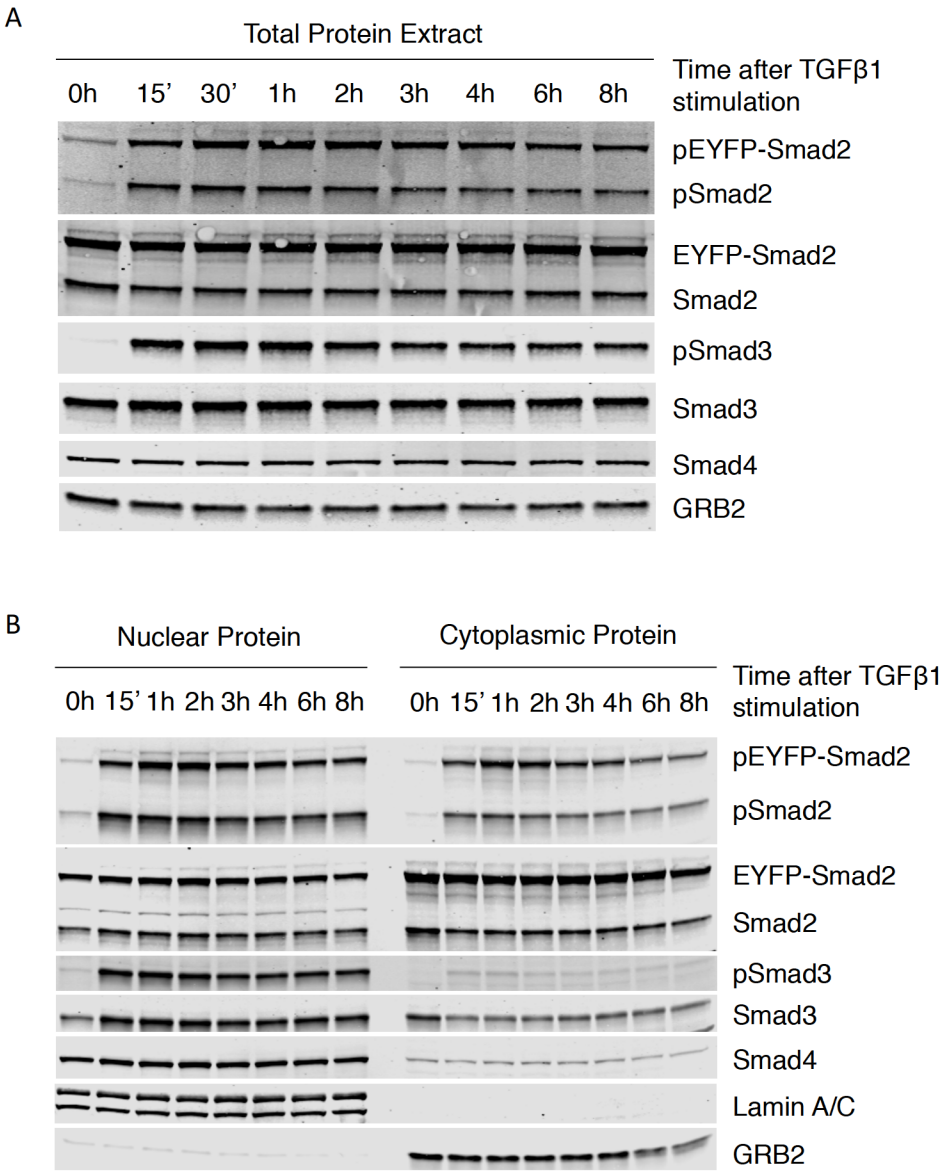


Figure 2.5 Western blot analysis of EYFP-Smad2 and endogenous Smad proteins. (A) Time course of Smad proteins in response to 100pM TGFβ1 stimulation. (B) Nuclear and cytoplasmic protein extracts were analyzed by western blotting with antibodies as indicated. The cells were treated the same as in (A). Experiments were repeated twice independently with similar results.

Thus, the EYFP-Smad2 in the stable cell line could be used to faithfully report the TGFβ signaling activity.

Taken together, we generated a system to monitor and quantitatively analyze TGFβ/Smad signaling and cell cycle progression simultaneously in live cells.

2.3 Dose-dependent G1 arrest induced by TGFβ at the population level

Previous studies showed that when treated with TGFβ, the fraction of cells in G1 phase increased[56], or sustained at a high level after being released from growth arrest[61].

To study if the TGFβ-induced G1 arrest was dose-dependent and phase-dependent, we decided to treat the cells with various doses of TGFβ (0, 5, 10, 20, 50, 100 or 200pM) for 3h and monitored the fraction of cells in G1 phase under the microscope over 60h (Figure 2.6).

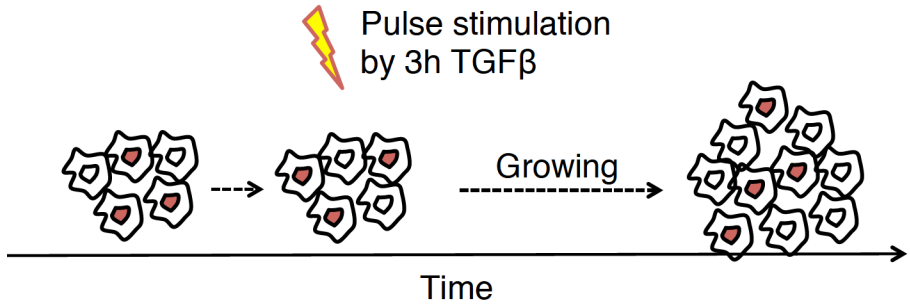


Figure 2.6 Schematic showing the treatment in experiments. The growing cells were monitored under the microscope over 60h. In the first 3h, the cells were treated with varied doses of TGFβ. After that, the TGFβ was removed.

The stimulation time was set as 3h for two main reasons:

Firstly, we were curious about the cell cycle phase dependent responses induced by TGFβ. Since the cells are growing when treated with TGFβ, some of them progress form one cell cycle phase to the next during the treatment (Figure 2.7A). The fraction of these cells increases when the stimulation duration is lengthened (theoretically analysis by mathematical modeling[133], Figure 2.7B). The average G1, S/G2/M phase duration is 8h, 12h, respectively. The time window of 3h treatment largely avoids the stimulation over two phases and facilitates analyzing cell cycle phase dependent responses.

Secondly, the nuclear accumulation of Smad2 reaches the maximum in 2h after stimulation (Figure 2.4A and B and Figure 2.5B). The 3h imaging could capture the amplitude of nuclear Smad2 dynamics and avoid measurement aberrations, like photo beaching.

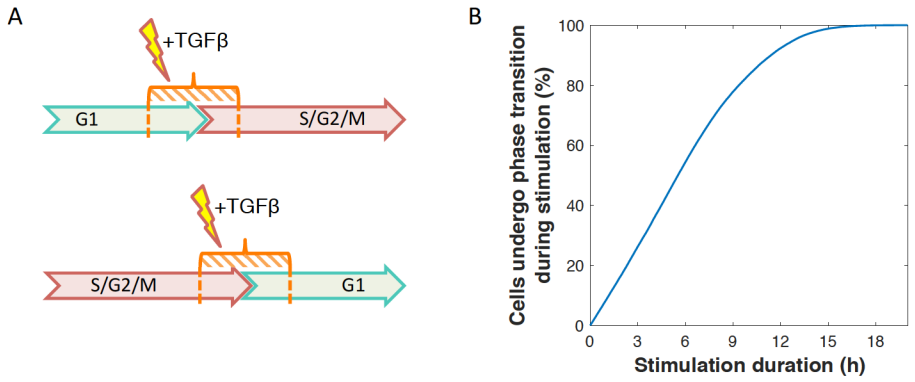


Figure 2.7 Cells undergo phase transition during TGFβ stimulation. (A) Schematic showing the cells undergo G1/S phase transition or M/G1 phase transition during the TGFβ treatment. (B) Theoretically analysis indicates that the fraction of cells undergo phase transition during stimulation increased with lengthened stimulation durations.

After data acquisition, we first checked the population behavior in response to the various doses of TGFβ. According to the knowledge in literature, we expected a larger fraction of cells in G1 phase after TGFβ stimulation. The number of cells with or without mCherry signal (in S/G2/M or G1 phase, respectively) was counted in each images and the fraction of cells in G1 phase (G1 fraction) was calculated and plotted (Figure 2.8).

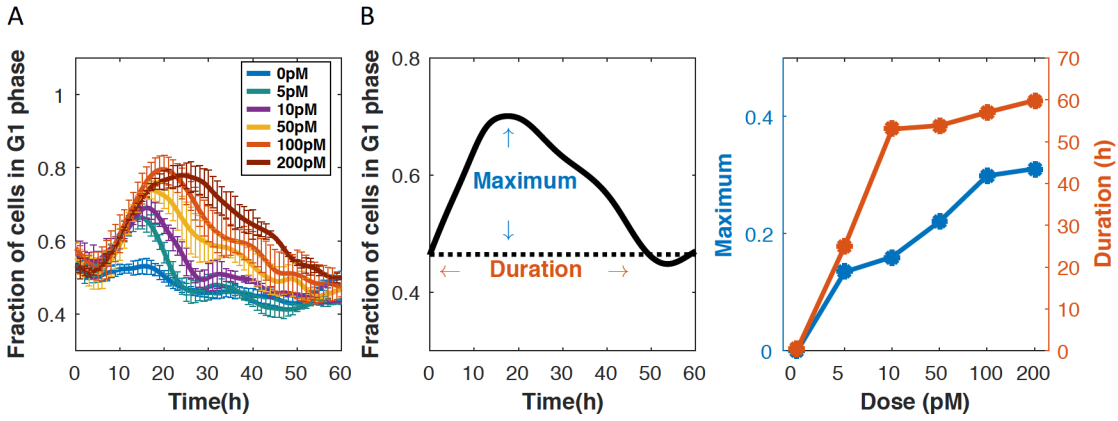


Figure 2.8 G1 arrest induced by TGFβ at population level. (A) The fractions of the cells in G1 phase were monitored after 3h of TGFβ1 stimulation (0,5,10,50,100 or 200pM). TGFβ was added at 0h and removed at 3h. The experiments were repeated at least 3 times independently. The data was obtained by calculating the average values of experimental data from 3 to 7 replicates. Error bars: stand error of the mean. (B) Left: Schematic definitions of Maximum and Duration; Maximum describes the largest difference between the G1 fraction of experimental group and that of control group; Duration describes the time period, during which the G1 fraction of experimental group is larger than control. Right: The maximum and the duration increase upon different concentrations of TGFβ treatment in (A).

Around 3h after TGFβ treatment, the fraction of cells in G1 phase increased, which reached the maximum in 15-20h and then fell back to the initial level. These

results were consistent with the previous observations of reversible G1 arrest induced by TGFβ[56,61]. Increasing TGFβ1 concentration elevated the maximum and duration of the response (Figure 2.8B). Notably, although the duration was dramatically elevated from 0pM to 10pM, the durations were similar from 10pM to 200pM. The maximum did not increase dramatically over 100pM.

The response mode of cell cycle arrest in individual cells was not clear. Based on these observations (Figure 2.8) at population level, we proposed two hypotheses of how the cell response to increasing dose of stimulus: graded and all-or-none response (Figure 2.9):

1) The graded response: all cells are arrested in G1 phase. A graded slow-down of progression in G1 phase is triggered by increased doses of TGFβ and the average arrested time is adjusted according to the concentration.
2) The all-or-none response: partial cells are arrested in G1 phase. The affected fraction is proportional to the ligand dose.

Both of these two different responses lead to a larger fraction of cells in G1phase on population average. To figure out which hypothesis is correct, we need to know the trajectories of individual cells upon TGFβ stimulation and analyze their G1 phase durations.

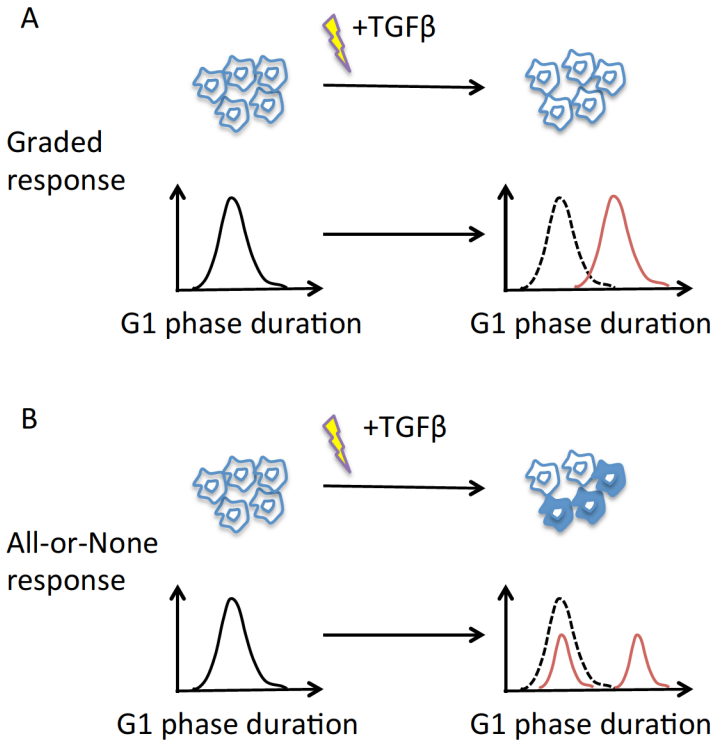


Figure 2.9 Hypotheses of cell arrest in response to TGFβ at single cell level. The response could be in Graded (A) or All-or-none (B) manners.

We tried to use some tools for automatic single cell lineage tracking, e.g. Cell profiler. However, it turned out that the accuracy was unacceptably low for long-term live cell imaging analysis. Some characteristics of the growing HaCaT cells leaded to the failure in lineage tracking, such as high mobility, overlapping, cell division and heterogeneous nuclear marker signals. Introducing manual correction could increase the accuracy but was extremely time-consuming.

Since computational or mathematical modeling is very helpful in prediction and gives guides for designing experiments, we decided to test the hypotheses *in silico* by mathematical modeling as the first step.

Chapter 3

Mathematical modeling of cell cycle regulation by TGFβ

In this chapter, we introduce how the population balance model of cell cycle (PBM) was developed, which could describe the numbers of growing cells over time.

The effects of TGFβ on cell cycle were simulated based on the assumption of either graded response or all-or-none response. Computational simulations suggested inheritable G1 arrest after cell division and strongly supported the hypothesis that TGFβ induces all-or-none G1 arrest response.

3.1 The population balance model of cell cycle (PBM)

Mechanistic models (containing protein interaction networks) of the cell cycle have been extensively developed to explain the molecular mechanism of cell growth. Compared with them, descriptive models focus on the measurable cause-effect phenomena to describe the overall system behavior while avoiding modeling of complex molecular mechanism[134].

Because our primary focus in this work was the G1 arrest phenomena in response to TGFβ treatment, we decided to use a descriptive model. Here we developed a population balance model of cell cycle (PBM) to describe the kinetics of cell population in each phase[135].

By applying the PBM to simulate cells in response to TGFβ (graded or all-or-none response mode), we could predict the fluctuation of the G1 fraction (the fraction of cells in G1 phase) over time. By comparing the model predictions with experimental data, we could find out the response mode supported by the model simulation and then verify it by experiments.

3.1.1 Model framework

To make the conception of PBM straightforward and easy to understand, we will introduce three models of cell cycle in the section, from the simplest CCP-ODE model to PBM (Table 3-1, Figure 3.1), showing how the PBM was evolved from the simplest CCP-ODE model.

Table 3-1 Models of cell cycle progression

	Models*	Phase-to-phase progression	Intra-phase information	Heterogeneous phase durations	Reference
1	CCP-ODE	✓	-	-	134,136
2	CCP-IntraPhase	✓	✓	-	135
3	PBM	✓	✓	✓	134,135

*CCP: cell cycle phase; ODE: ordinary differential equations. ✓: the model describes the characteristic.

3.1.1.1 CCP-ODE model

To describe the numbers of cells in different cell cycle phases, the simplest approach was to use two variables (G and M) to describe the number of cells in G1 and S/G2/M phases (Figure 3.1A). In this simple model, there were two parameters (τ_G and τ_M) to represent each phase durations, and two ordinary differential equations (ODEs) to describe phase transitions (Equation 1 and 2):

$$\frac{\mathrm{dG}}{\mathrm{dt}} = 2\frac{M}{\tau_M} - \frac{G}{\tau_G} \tag{1}$$

$$\frac{\mathrm{dM}}{\mathrm{dt}} = \frac{G}{\tau_G} - \frac{M}{\tau_M} \tag{2}$$

where G and M are the cell numbers in G1 and S/G2/M phases, respectively; τ_G and τ_M are their respective phase durations.

3.1.1.2 CCP-IntraPhase model

However, this CCP-ODE model only described the averaged phase behavior and neglects the intraphase coordinates. Therefore a revised model was developed, considering the intra-phase information (CCP-IntraPhase model). In the model, we divided G1 and S/G2/M phases into small sub-phases, cells in each sub-phase could move to the next sub-phase (Figure 3.1B).

We could describe the number of cells in each sub-phase by one ordinary differential equation (Equation 3 and 4):

$$\frac{\mathrm{d}G_i}{\mathrm{dt}} = \begin{cases} 2\dfrac{M_m}{\tau} - \dfrac{G_i}{\tau} & for\ i = 1 \\ \dfrac{G_{i-1}}{\tau} - \dfrac{G_i}{\tau} & for\ i = 2, \dots, n \end{cases} \tag{3}$$

$$\frac{\mathrm{d}M_j}{\mathrm{dt}} = \begin{cases} \dfrac{G_n}{\tau} - \dfrac{M_j}{\tau} & for\ j = 1 \\ \dfrac{M_{j-1}}{\tau} - \dfrac{M_j}{\tau} & for\ j = 2, \dots, m \end{cases} \tag{4}$$

where G_i and M_j are the cell numbers in the i-th sub-phase in G1 and j-th sub-phase in S/G2/M, respectively; τ is the duration of the sub-phase.

Notably, here we set the same length (τ) for all the sub-phases in the cell cycle to avoid introducing noise. Thus the G1 and S/G2/M phases were divided to

certain numbers of sub-phases (here n and m for G1 and S/G2/M phases, respectively) according to the maximum length of the respective phase duration measured in experiments. Moreover, the finer sub-phase we divide, the larger numbers of the ODEs, and the better it describes the cell cycle progression.

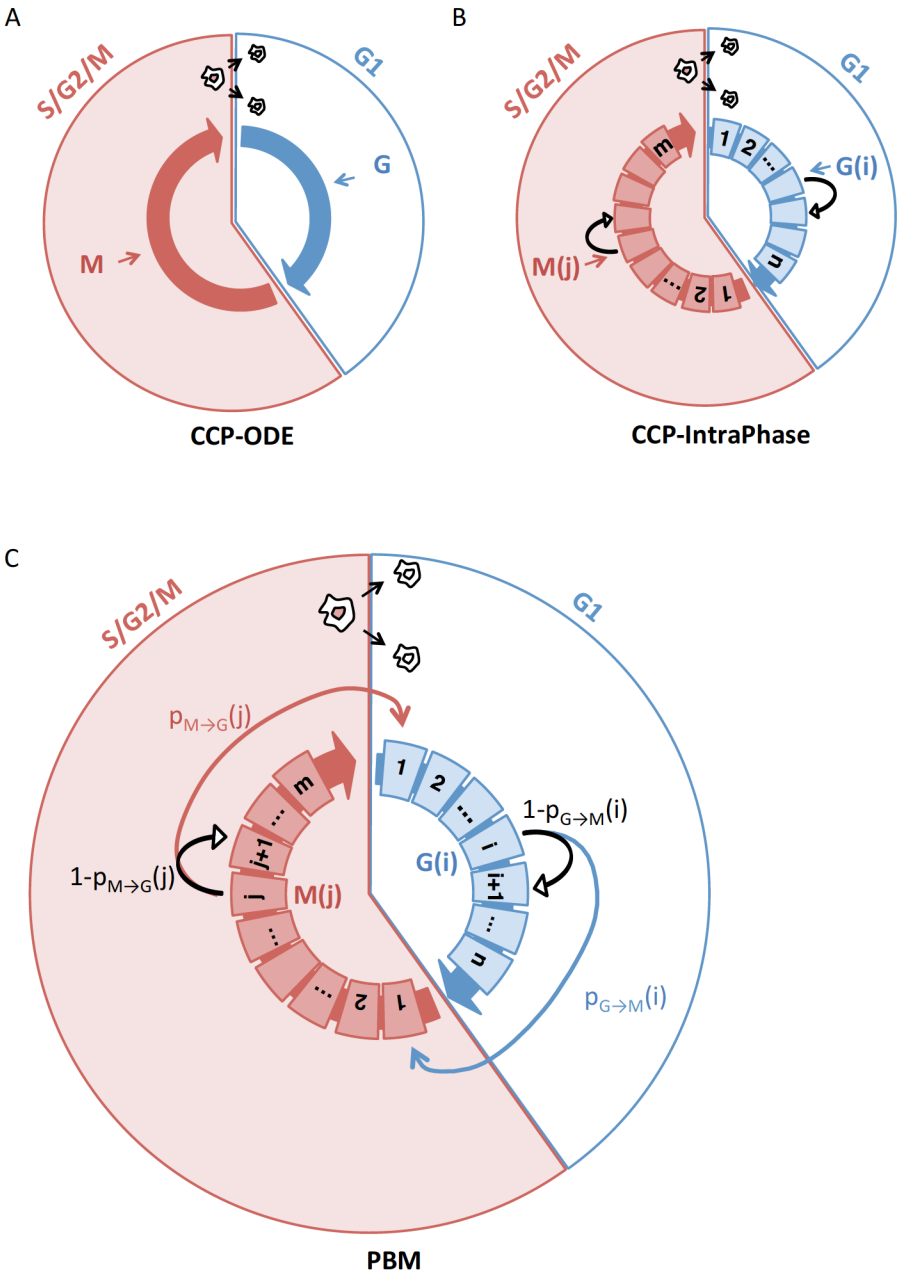


Figure 3.1 Schematic diagrams for model conceptions. (A) CCP-ODE model: a simple model with two variables describing the number of cells in G1 and S/G2/M phases, and two ODEs to describe phase transitions. (B) CCP-IntraPhase model: a revised model by considering the intra-phase information. In this model, G1 and S/G2/M phases are divided into fine sub-phases. The cells in each sub-phase can move to the next sub-phase. (C) PBM: the population balance model, considering the intra-phase information and noise (Varied phase durations in individual cells). The cells in each sub-phase can either move to the next sub-phase, or directly move to the first sub-phase of the next cell-cycle phase.

3.1.1.3 PBM (population balance model of cell cycle)

The CCP-IntraPhase model assumed that every cell has the same G1 length ($n \cdot \tau$) and same S/G2/M length ($m \cdot \tau$). However, noise exists in reality to vary the cell cycle phase durations in the cell population. When reflected in the model, a cell in each sub-phase could not only move to the next sub-phase, but could also directly transit to the first sub-phase of the next phase. Thus the PBM was developed, considering heterogeneous cell cycle phase durations (Figure 3.1C).

To introduce noise (to vary the cell cycle phase durations) into the model, we introduced a set of parameters of phase transition probabilities (i.e. the probability of a cell transiting into next cell cycle phase at certain sub-phase): $p_{G\to M}$ or $p_{M\to G}$, for cells in G1 or S/G2/M phase, respectively. For the cells in each sub-phase, they have two possible fates: (1) transit to the first sub-phase of the next phase (with a probability $p_{G\to M}$ or $p_{M\to G}$); (2) move to the next sub-phase (with a probability $1-p_{G\to M}$ or $1-p_{M\to G}$). Therefore, we could describe each sub-phase by one ODE while taking noise into consideration (Equation 5 and 6):

$$\frac{\mathrm{d}G_i}{\mathrm{dt}} = \begin{cases} 2\frac{M_m}{\tau}\left(1-p_{M\to G}(m)\right) + 2\sum_{j=1}^{m} M_j \cdot p_{M\to G}(j) - \frac{G_i}{\tau}\left(1-p_{G\to M}(i)\right) - G_i \cdot p_{G\to M}(i) \\ \qquad\qquad for\ i = 1 \\ \frac{G_{i-1}}{\tau}\left(1-p_{G\to M}(i-1)\right) - \frac{G_i}{\tau}\left(1-p_{G\to M}(i)\right) - G_i \cdot p_{G\to M}(i) \\ \qquad\qquad for\ i = 2, \dots, n \end{cases} \tag{5}$$

$$\frac{\mathrm{d}M_j}{\mathrm{dt}} = \begin{cases} \frac{G_n}{\tau}\left(1-p_{G\to M}(n)\right) + \sum_{i=1}^{n} G_i \cdot p_{G\to M}(i) - \frac{M_j}{\tau}\left(1-p_{M\to G}(j)\right) - M_j \cdot p_{M\to G}(j) \\ \qquad\qquad for\ j = 1 \\ \frac{M_{j-1}}{\tau}\left(1-p_{M\to G}(j-1)\right) - \frac{M_j}{\tau}\left(1-p_{M\to G}(j)\right) - M_j \cdot p_{M\to G}(j) \\ \qquad\qquad for\ j = 2, \dots, m \end{cases} \tag{6}$$

where G_i and M_j are the cell numbers in the i-th sub-phase in G1 and j-th sub-phase in S/G2/M, respectively; τ is the duration of the sub-phase; $p_{G\to M}(i)$ and $p_{M\to G}(j)$ are the conditional phase transition probabilities of the i–th sub-phase in G1 and the j–th sub-phase in S/G2/M, respectively.

Up to now, we developed a PBM containing two sets of ODEs to describe the cell cycle progressions of heterogeneous proliferating cells.

3.1.2 Model assumptions

The PBM was developed based on the following assumptions:

(1) The cell cycle is consisted of four phases: G1, S, G2 and M phase. As we focus on G1 phase, we divide cell cycle duration into two parts: G1 phase duration and S/G2/M phase duration. Each phase duration (G1 or S/G2/M phase) is assumed to follow a lognormal distribution[137,138].

(2) One cell divides into two cells after the completion of the S/G2/M phase.

(3) We observed few cells died in our experiments. To simplify the model, the cell death is not considered here in the model. But the PBM can be extended to include cell death in the future studies.

3.1.3 Estimating the parameters of the PBM

Next, we estimated the values of parameters in the PBM (Table 3-2).

Table 3-2 Summary of the parameters in PBM

	Parameter	Annotation
1	μ_G	Parameter μ of the log-normally distributed G1 phase duration
2	σ_G	Parameter σ of the log-normally distributed G1 phase duration
3	μ_M	Parameter μ of the log-normally distributed S/G2/M phase duration
4	σ_M	Parameter σ of the log-normally distributed S/G2/M phase duration
5	τ	The duration of the subphase
6	n	The number of subphases in G1 phase
7	m	The number of subphases in S/G2/M phase
8	$p_{G\to M}(i)$	The probability of cells transiting into S/G2/M phase at i-th sub-phase in G1 phase; i=1, … , n
9	$p_{M\to G}(j)$	The probability of cells transiting into G1 phase at j-th sub-phase in S/G2/M phase; j=1, … , m

We quantitatively measured the G1 or S/G2/M phase durations of untreated cells in experiments. The data fitted to lognormal distribution functions and the values of parameters (μ_G, σ_G, μ_M and σ_M) were calculated (Figure 3.2).

After that, we set the value of τ, n and m. The theoretical 99% confidence intervals of G1 and S/G2/M phase durations were calculated base on the values of μ_G, σ_G, μ_M and σ_M. In the model, we divided G1 and S/G2/M phases into n and m sub-phases. The durations of sub-phases were the same (τ). The length of G1 ($n \cdot \tau$) and S/G2/M ($m \cdot \tau$) phase were set as the upper endpoints of the confidence intervals. The finer sub-phase we divide, the better it describes the cell cycle progression and the larger numbers of the ODEs in the model. We found that the model with n larger than 500 was acceptable to describes the cell cycle progression of our cell line used in experiments. We decided to divide G1 phase into 550 sub-phases (n=550). Then the values of parameters τ and m were calculated accordingly.

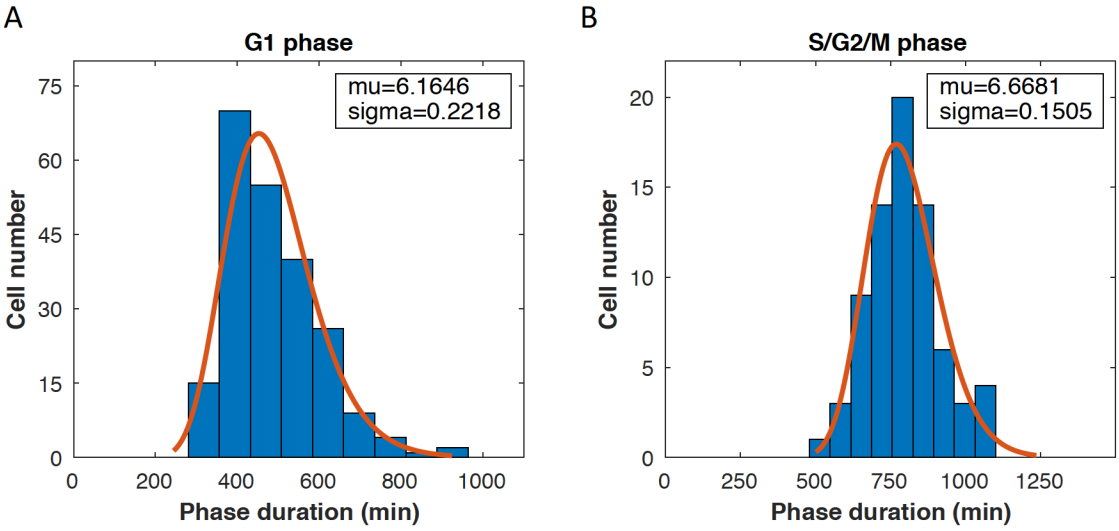


Figure 3.2 Distributions of cell cycle phase durations. (A) G1 phase duration distribution. The G1 phase durations of 222 cells experimentally measured. (B) S/G2/M phase duration distribution. The S/G2/M phase durations of 74 cells were experimentally measured. The experimental data fits to lognormal density functions and the values of parameters (mu and sigma) were calculated. Those cells were randomly chosen from three independently repeated experiments.

Knowing the values of τ, n, m, μ_G, σ_G, μ_M and σ_M, the transition probability ($P_{G\to M}$ or $P_{M\to G}$) could be calculated. Lognormal possibility density functions of G1 or S/G2/M phase durations were simulated (Figure 3.3A, Equation 7 and 8).

$$P_{G1\ duration} = \sim lognomalPDF(\mu_G, \sigma_G) \quad (7)$$

$$P_{S/G2/M\ duration} = \sim lognomalPDF(\mu_M, \sigma_M) \quad (8)$$

Notably, these distributions (Equation 7 and 8) represent the probabilities of phase durations, which are actually conditional probabilities (i.e. the probability of

cell transiting directly to the next phase under the condition that it did not transit in the previous sub-phases). For cells in each sub-phase, the phase transition probability ($P_{G\to M}$ or $P_{M\to G}$) could be calculated accordingly (Figure 3.3B, Equation 9 and 10).

$$p_{G1\ duration}(i) = p_{G\to M}(i) * (1 - p_{G\to M}(i-1)) * \dots * (1 - p_{G\to M}(1)) \qquad i = 1, \dots, n \qquad (9)$$

$$p_{S/G2/M\ duration}(j) = p_{M\to G}(j) * (1 - p_{M\to G}(j-1)) * \dots * (1 - p_{M\to G}(1)) \qquad j = 1, \dots, m \qquad (10)$$

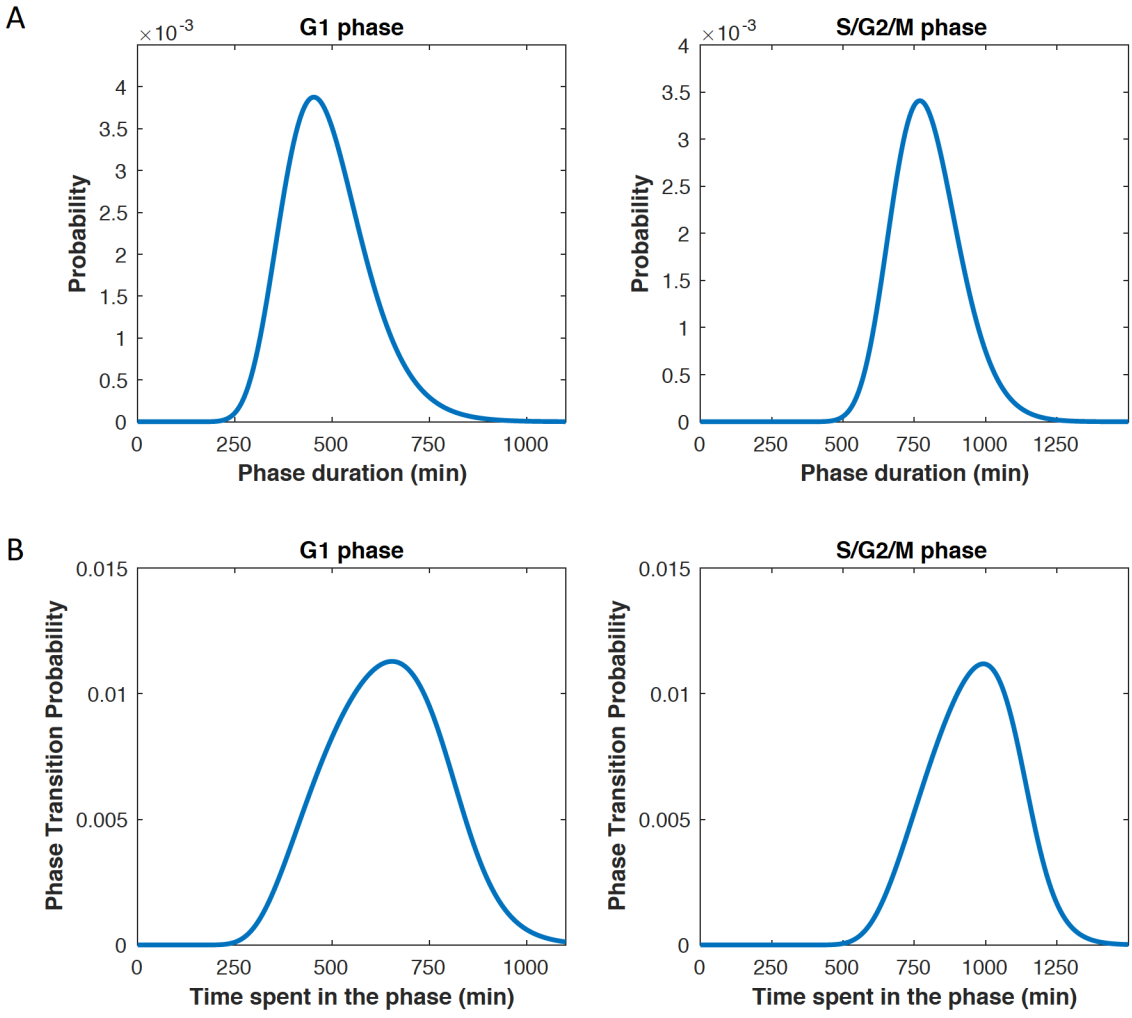


Figure 3.3 Examples showing the method to calculate phase transition probabilities. (A) Lognormal possibility density functions of G1 or S/G2/M phase durations were calculated based on the values of μ_G, σ_G, μ_M and σ_M in Table 3-3 (C) The phase transition probabilities were calculated based on the data in (B).

The values of the parameters in the PBM were estimated (Table 3-3).

Table 3-3 The values of parameters in PBM

	Parameter	Value and Unit
1	μ_G	6.1646 min
2	σ_G	0.2218 min
3	μ_M	6.6681 min
4	σ_M	0.1505 min
5	τ	2.6213 min
6	n	550
7	m	637
8	$p_{G\to M}(i)$	(See Appendix Chapter 9: 9.1)
9	$p_{M\to G}(j)$	(See Appendix Chapter 9: 9.1)

3.1.4 Steady-state distribution for initial condition

We assumed that the population has reached the steady state before TGFβ treatment. At the steady state, the fractions of the asynchronous cells in each sub-phase are constant[133]. In theory, we could perform steady state analysis to derive the percentage of cells in each sub-phase. We rewrote equation 1 and 2 in terms of percentage:

$$\mathrm{G}_i^{\%} = \frac{G_i}{A} \tag{11}$$

$$\mathrm{M}_j^{\%} = \frac{M_j}{A} \tag{12}$$

where $\mathrm{G}_i^{\%}$ and $\mathrm{M}_j^{\%}$ are percentage of cells in the i-th sub-phase in G1 and j-th sub-phase in S/G2/M, respectively; $A = \sum_{i=1}^{n} G_i + \sum_{j=1}^{m} M_j$. Therefore,

$$\frac{\mathrm{dG}_i^{\%}}{\mathrm{dt}} = \frac{\mathrm{d}G_i}{\mathrm{dt}} \cdot \frac{1}{A} - \mathrm{G}_i^{\%} \cdot \frac{\mathrm{dA}}{\mathrm{dt}} \cdot \frac{1}{A} \tag{13}$$

$$\frac{\mathrm{dM}_j^{\%}}{\mathrm{dt}} = \frac{\mathrm{d}M_j}{\mathrm{dt}} \cdot \frac{1}{A} - \mathrm{M}_j^{\%} \cdot \frac{\mathrm{dA}}{\mathrm{dt}} \cdot \frac{1}{A} \tag{14}$$

$$\frac{\mathrm{dA}}{\mathrm{dt}} \cdot \frac{1}{A} = \frac{M_m^{\%}}{\tau}\left(1 - p_{M\to G}(m)\right) + \sum_{j=1}^{m} M_j^{\%} \cdot p_{M\to G}(j) \tag{15}$$

By replacing $\frac{dG_i}{dt}$, $\frac{dM_j}{dt}$, A and $\frac{dA}{dt}\cdot\frac{1}{A}$ with the equation 5, 6 and 15. Then equations 13 and 14 could be transformed to:

$$\frac{dG_i^{\%}}{dt}=\begin{cases} 2\frac{M_m^{\%}}{\tau}\left(1-p_{M\to G}(m)\right)+2\sum_{j=1}^{m}M_j^{\%}\cdot p_{M\to G}(j)-\frac{G_i^{\%}}{\tau}\left(1-p_{G\to M}(i)\right)-G_i^{\%}\cdot p_{G\to M}(i) \\ -G_i^{\%}\cdot\left(\frac{M_m^{\%}}{\tau}\left(1-p_{M\to G}(m)\right)+\sum_{j=1}^{m}M_j^{\%}\cdot p_{M\to G}(j)\right) \qquad for\ i=1 \\ \frac{G_{i-1}^{\%}}{\tau}\left(1-p_{G\to M}(i-1)\right)-\frac{G_i^{\%}}{\tau}\left(1-p_{G\to M}(i)\right)-G_i^{\%}\cdot p_{G\to M}(i) \\ -G_i^{\%}\cdot\left(\frac{M_m^{\%}}{\tau}\left(1-p_{M\to G}(m)\right)+\sum_{j=1}^{m}M_j^{\%}\cdot p_{M\to G}(j)\right) \qquad for\ i=2,\dots,n \end{cases} \tag{16}$$

$$\frac{dM_j^{\%}}{dt}=\begin{cases} \frac{G_n^{\%}}{\tau}\left(1-p_{G\to M}(n)\right)+\sum_{i=1}^{n}G_i^{\%}\cdot p_{G\to M}(i)-\frac{M_j^{\%}}{\tau}\left(1-p_{M\to G}(j)\right)-M_j^{\%}\cdot p_{M\to G}(j) \\ -M_j^{\%}\cdot\left(\frac{M_m^{\%}}{\tau}\left(1-p_{M\to G}(m)\right)+\sum_{j=1}^{m}M_j^{\%}\cdot p_{M\to G}(j)\right) \qquad for\ j=1 \\ \frac{M_{j-1}^{\%}}{\tau}\left(1-p_{M\to G}(j-1)\right)-\frac{M_j^{\%}}{\tau}\left(1-p_{M\to G}(j)\right)-M_j^{\%}\cdot p_{M\to G}(j) \\ -M_j^{\%}\cdot\left(\frac{M_m^{\%}}{\tau}\left(1-p_{M\to G}(m)\right)+\sum_{j=1}^{m}M_j^{\%}\cdot p_{M\to G}(j)\right) \qquad for\ j=2,\dots,m \end{cases} \tag{17}$$

Each $G_i^{\%}$ and $M_j^{\%}$ at steady state could be derived by setting $\frac{dG_i^{\%}}{dt}=0$ and $\frac{dM_j^{\%}}{dt}=0$, and then solving the systems of equations.

The initial distribution of the cells in each subphase does not affect the steady-state distribution. For convenience, here we generated the quasi steady state by setting random values for cell number of each sub-phases (for simplicity, we set G_i = M_j = 10 cells) and running a long-term simulation until the quasi steady state was reached. We then stored it as the initial condition (G^{steady} and M^{steady}) for the simulations later on. The initial conditions of the variables are summarized (Table 3-4).

Table 3-4 Summary of model initial conditions

Variable	Annotation	Initial Conditions
G_i	The cell numbers in the i-th sub-phase of G1 phase	G_i^{steady}
M_j	The cell numbers in the j-th sub-phase of S/G2/M phase	M_j^{steady}

3.2 Computational modeling of TGFβ-induced G1 arrest

3.2.1 The PBM incorporating the effect of ligand-induced G1 arrest

In order to study how TGFβ regulates G1 arrest, we next incorporated the effect of ligand-induced G1 arrest into the PBM.

Notably, asynchronized cells are either in G1 phase or S/G2/M phase at the timing of TGFβ stimulation. TGFβ primarily targets G1 events[9] and cells in G1 phase are arrested[61]. It was unclear how the cells in S/G2/M phase respond to TGFβ. We proposed two assumptions here (Figure 3.4): Assumption 1, cells in G1 phase at the time of TGFβ stimulation are arrested, while the cells in S/G2/M phase are unaffected (Figure 3.4A); Assumption 2, upon the TGFβ stimulation, the cells in G1 phase are arrested immediately (1st G1 phase) and the cells in S/G2/M phase will be arrested in the G1 phase of the next cell cycle (2nd G1 phase)(Figure 3.4B).

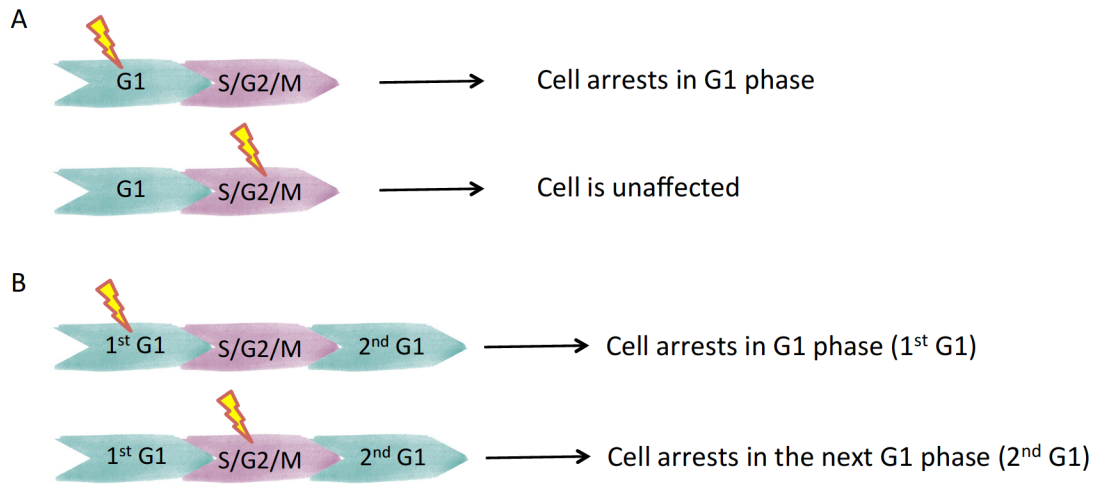


Figure 3.4 Assumptions of modeling TGFβ-induced G1 arrest. Upon the ligand stimulation, (A) the cells in G1 phase are arrested, while the cells in S/G2/M phase are unaffected; (B) the cells in G1 phase are arrested immediately (1st G1 phase) and the cells in S/G2/M phase will be arrested in the G1 phase of the next cell cycle (2nd G1 phase).

We revised the PBM to make it applicable to represent ligand-induced G1 arrest of either assumption. In the model, after ligand stimulation, cells could be unaffected or affected. We employed two parameter p'_G and p'_M to describe the probability of cells in G1 or S/G2/M phase being affected (For assumption 1, p'_M =0).

For cells in G1 phase, a new group of arrested cells is generated by applying p'_G to the cells in all the sub-phases before stimulation: $G' = G \cdot p'_G$. The rest stay in unaffected group: $G = G \cdot (1 - p'_G)$. For arrested cells, they have different G1 phase duration compared with the untreated cells. Therefore, we set a new parameter n' for the number of sub-phases in arrested G1 phase. The arrested G1 duration equals to $n' \cdot \tau$. The phase transition probability of arrested cells is $P'_{G \to M}$, which can be calculated in a similar way as in Chapter 3: 3.1.3 (Figure 3.5).

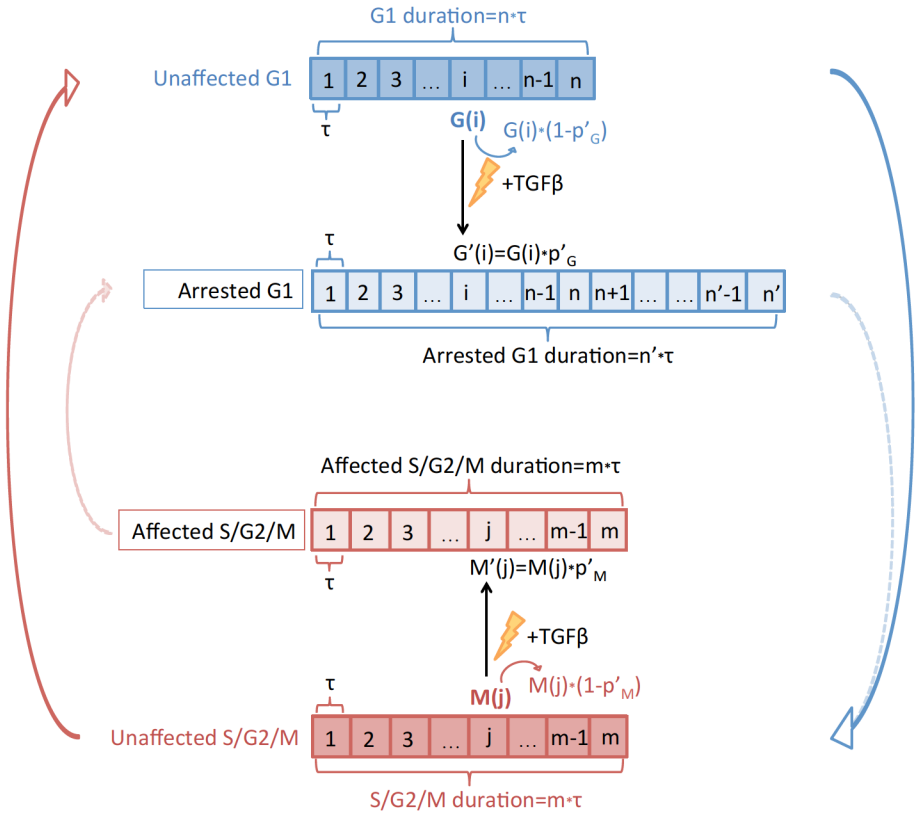


Figure 3.5 Schematic diagrams of PBM incorporating the effect of ligand-induced G1 arrest. After TGFβ stimulation, cells in G1 phase have two fates: be arrested (with the probability p'_G), unaffected (with the probability 1-p'_G); for cells in each sub-phase, partial cells ($G(i) \cdot p'_G$) move to arrested group, the rest stay in unaffected group. Cells in S/G2/M phase have two fates: be affected (with the probability p'_M), unaffected (with the probability 1-p'_M); for cells in each sub-phase, partial cells ($M(j) \cdot p'_M$) move to affected group, the rest stay in unaffected group. Unaffected G1, arrested G1, S/G2/M and affected S/G2/M phase are divided into n, n', m and m' sub-phases. τ : the duration of the sub-phase; G(i): cell numbers in i-th sub-phase in G1 phase; G'(i): cell numbers in i-th sub-phase in arrested G1 phase; M(j): cell numbers in j-th sub-phase in S/G2/M phase. M'(j): cell numbers in j-th sub-phase in affected S/G2/M phase.

For cells in S/G2/M phase, a new group of affected cells is generated by applying p'_M to the cells in all the sub-phases before stimulation: $M' = M \cdot p'_M$. The rest stay in unaffected group: $M = M \cdot (1 - p'_M)$. For affected cells, their S/G2/M phase durations and phase transition probabilities are not changed, but they will move to the G1 arrested group G'(Figure 3.5).

We could similarly describe each sub-phase by one ODE for the affected cells (Equation 18 and 19):

$$\frac{dG'_i}{dt} = \begin{cases} 2\frac{M'_m}{\tau}\left(1 - p_{M\to G}(m)\right) + 2\sum_{j=1}^{m} M'_j \cdot p_{M\to G}(j) - \frac{G'_i}{\tau}\left(1 - p'_{G\to M}(i)\right) - G'_i \cdot p'_{G\to M}(i) & \\ & for\ i = 1 \\ \frac{G'_{i-1}}{\tau}\left(1 - p'_{G\to M}(i-1)\right) - \frac{G'_i}{\tau}\left(1 - p'_{G\to M}(i)\right) - G'_i \cdot p'_{G\to M}(i) & \\ & for\ i = 2, \dots, n \end{cases} \tag{18}$$

$$\frac{dM'_j}{dt} = \begin{cases} -\frac{M'_j}{\tau}\left(1 - p_{M\to G}(j)\right) - M'_j \cdot p_{M\to G}(j) & for\ j = 1 \\ \frac{M'_{j-1}}{\tau}\left(1 - p_{M\to G}(j-1)\right) - \frac{M'_j}{\tau}\left(1 - p_{M\to G}(j)\right) - M'_j \cdot p_{M\to G}(j) & for\ j = 2, \dots, m \end{cases} \tag{19}$$

where G'_i and M'_j are the cell numbers of affected cells in the i-th sub-phase in G1 and j-th sub-phase in S/G2/M phase; τ is the duration of the sub-phase, which is the same as it in unaffected group; $p'_{G\to M}(i)$ is the phase transition probabilities of the i–th sub-phase in arrested G1 phase, could be calculated in the similar way as $p_{G\to M}(i)$. $p_{M\to G}(j)$ is the phase transition probabilities of the i–th sub-phase in affected S/G2/M phase, which is the same as it in unaffected group.

The ODEs describing cell numbers in unaffected S/G2/M phase (Equation 6) were modified to include the cells that were arrested in G1 phase (Equation 20).

$$\frac{dM_j}{dt} = \begin{cases} \frac{G_n}{\tau}\left(1 - p_{G\to M}(n)\right) + \sum_{i=1}^{n} G_i \cdot p_{G\to M}(i) - \frac{M_j}{\tau}\left(1 - p_{M\to G}(j)\right) - M_j \cdot p_{M\to G}(j) & \\ + \frac{G'_{n'}}{\tau}\left(1 - p'_{G\to M}(n')\right) + \sum_{i=1}^{n'} G'_i \cdot p'_{G\to M}(i) & for\ j = 1 \\ \frac{M_{j-1}}{\tau}\left(1 - p_{M\to G}(j-1)\right) - \frac{M_j}{\tau}\left(1 - p_{M\to G}(j)\right) - M_j \cdot p_{M\to G}(j) & for\ j = 2, \dots, m \end{cases} \tag{20}$$

Up to now, we developed a PBM that could describe cell cycle dynamics upon ligand stimulation. This model is consisted of two parts: one describes the normal cell cycle progression; the other describes cells that are arrested by the ligand. They are linked by the parameters p'_G and p'_M, which describe the probability of cells in G1 or S/G2/M phase being affected upon the ligand stimulation.

The equations in the PBM are summarized (Table 3-5).

Table 3-5 Summary of the equations in PBM incorporating the ligand-induced G1 arrest

Untreated and unaffected cells

$$\frac{dG_i}{dt} = \begin{cases} 2\frac{M_m}{\tau}\left(1 - p_{M\to G}(m)\right) + 2\sum_{j=1}^{m} M_j \cdot p_{M\to G}(j) - \frac{G_i}{\tau}\left(1 - p_{G\to M}(i)\right) - G_i \cdot p_{G\to M}(i) & for\ i = 1 \\ \frac{G_{i-1}}{\tau}\left(1 - p_{G\to M}(i-1)\right) - \frac{G_i}{\tau}\left(1 - p_{G\to M}(i)\right) - G_i \cdot p_{G\to M}(i) & for\ i = 2, \dots, n \end{cases}$$

$$\frac{dM_j}{dt} = \begin{cases} \frac{G_n}{\tau}\left(1 - p_{G\to M}(n)\right) + \sum_{i=1}^{n} G_i \cdot p_{G\to M}(i) - \frac{M_j}{\tau}\left(1 - p_{M\to G}(j)\right) - M_j \cdot p_{M\to G}(j) \\ \quad + \frac{G'_{n'}}{\tau}\left(1 - p'_{G\to M}(n')\right) + \sum_{i=1}^{n'} G'_i \cdot p'_{G\to M}(i) & for\ j = 1 \\ \frac{M_{j-1}}{\tau}\left(1 - p_{M\to G}(j-1)\right) - \frac{M_j}{\tau}\left(1 - p_{M\to G}(j)\right) - M_j \cdot p_{M\to G}(j) & for\ j = 2, \dots, m \end{cases}$$

Affected cells

$$\frac{dG'_i}{dt} = \begin{cases} 2\frac{M'_m}{\tau}\left(1 - p_{M\to G}(m)\right) + 2\sum_{j=1}^{m} M'_j \cdot p_{M\to G}(j) - \frac{G'_i}{\tau}\left(1 - p'_{G\to M}(i)\right) - G'_i \cdot p'_{G\to M}(i) & for\ i = 1 \\ \frac{G'_{i-1}}{\tau}\left(1 - p'_{G\to M}(i-1)\right) - \frac{G'_i}{\tau}\left(1 - p'_{G\to M}(i)\right) - G'_i \cdot p'_{G\to M}(i) & for\ i = 2, \dots, n \end{cases}$$

$$\frac{dM'_j}{dt} = \begin{cases} -\frac{M'_j}{\tau}\left(1 - p_{M\to G}(j)\right) - M'_j \cdot p_{M\to G}(j) & for\ j = 1 \\ \frac{M'_{j-1}}{\tau}\left(1 - p_{M\to G}(j-1)\right) - \frac{M'_j}{\tau}\left(1 - p_{M\to G}(j)\right) - M'_j \cdot p_{M\to G}(j) & for\ j = 2, \dots, m \end{cases}$$

Upon TGFβ stimulation, cells were categorized into four groups by applying the parameters p'_G and p'_M to the number of cells in each sub-phases (G_i and M_j) at quasi steady state. The initial conditions of the variables are summarized (Table 3-6).

Table 3-6 Initial conditions of PBM incorporating the ligand-induced G1 arrest

Variable	Annotation	Initial Conditions
G_i	The cell numbers in the i-th sub-phase of G1 phase	$G_i^{steady} \cdot (1 - p'_G(i))$
M_j	The cell numbers in the j-th sub-phase of S/G2/M phase	$M_j^{steady} \cdot (1 - p'_M(j))$
G'_i	The cell numbers in the i-th sub-phase of arrested G1 phase	$G_i^{steady} \cdot p'_G(i)$
M'_j	The cell numbers in the j-th sub-phase of affected S/G2/M phase	$M_j^{steady} \cdot p'_M(j)$

3.2.2 Testing hypotheses by PBM model simulations

To test whether TGFβ-induced G1 arrest in single cells is graded or all-or-none response (Figure 2.9) and whether the cells in S/G2/M phase are affected or not (Figure 3.4), we simulated the four PBMs with different assumptions (Table 3-7).

Table 3-7 PBM with different assumptions

	Graded response	All-or-None response
Cells in S/G2/M are unaffected	Model 1	Model 2
Cells in S/G2/M are arrested in 2nd G1 phase	Model 3	Model 4

3.2.2.1 Model simulation suggests the cells in S/G2/M phase are affected

In this section, we simulated Model 1 and Model 2, based on the assumption that cells in S/G2/M phase are unaffected by the cytostatic effect of TGFβ ($p'_M = 0$).

Different initial conditions were used for simulations of graded or all-or-none response. For each response mode, we did 10 simulations to represent cells treated with a series increasing doses of TGFβ. The cells without treatment were simulated as control (mean of G1 phase durations D_G=8.1h, arithmetic standard

deviation STD_G= 1.8h, $p'_G = 0$. For the arrested group, the mean of G1 phase durations D'_G and arithmetic standard deviation STD'_G of arrested group were set the same as control).

In the Graded response groups, all cells were affected (p'_G=1). The average G1 phase duration of arrested cells was increasing (D'_G, from 12.3 to 50h; the values were obtained from the arithmetic sequence from D_G = 8.1h to 50h).

In the all-or-none response groups, partial cells were affected. The affected fractions were increased (p'_G, from 0.1 to 1). The arrested cells were drawn from the same distribution (D'_G=40.5h, STD'_G =4.5h. The value of $D'_G = 5 * D_G$ and $STD'_G = 2.5 * STD_G$).

Table 3-8 The values of parameters for arrested cells in Model 1 and Model2

	Model 1:Graded response			Model 2:All-or-none		
	p'_G	D'_G (h)	STD'_G (h)	p'_G	D'_G (h)	STD'_G (h)
1(control)	0	8.1	1.8	0	8.1	1.8
2	1	12.3	1.8	0.1	40.5	4.5
3	1	16.5	1.8	0.2	40.5	4.5
4	1	20.7	1.8	0.3	40.5	4.5
5	1	24.9	1.8	0.4	40.5	4.5
6	1	29.1	1.8	0.5	40.5	4.5
7	1	33.2	1.8	0.6	40.5	4.5
8	1	37.4	1.8	0.7	40.5	4.5
9	1	41.6	1.8	0.8	40.5	4.5
10	1	45.8	1.8	0.9	40.5	4.5
11	1	50	1.8	1	40.5	4.5

The fractions of cells in G1 phases over time and the initial distributions of G1 durations were shown (Figure 3.6).

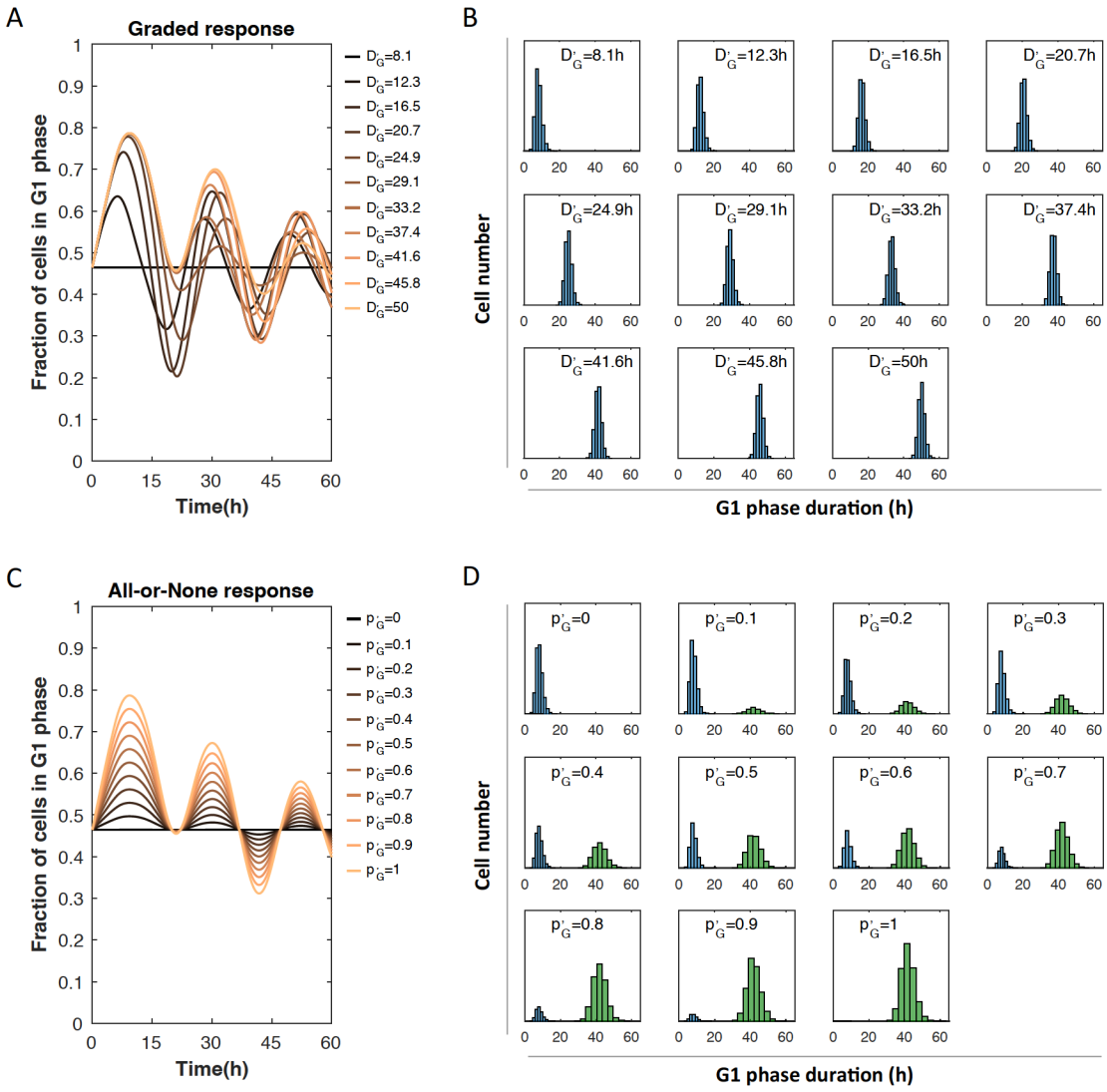


Figure 3.6 Comparison between simulation results of graded response and all-or-none response. (A) The simulation results of graded response: the fractions of cells in G1 phase over time. D'_G: The mean of arrested G1 phase durations (h). (B) The initial G1 phase duration distributions, corresponding to the curves in (A). (C) The simulation results of all-or-none response: the fractions of cells in G1 phase over time. p'_G: the probability of cells being arrested. (D) The initial G1 phase duration distributions, corresponding to the curves in (C); Blue distributions: unaffected cells; Green distributions: arrested cells.

The experimental data show that increasing TGFβ concentration elevated the maximum and duration of the responses (Figure 2.8). We next compared the model simulations with experimental data in terms of the maximums and durations. Compared with experimental data, of which the duration could reach 50-60h upon increasing doses of TGFβ treatment (Figure 3.7A), model simulations of both modes show durations that were no more than 20h (Figure 3.7B and C).

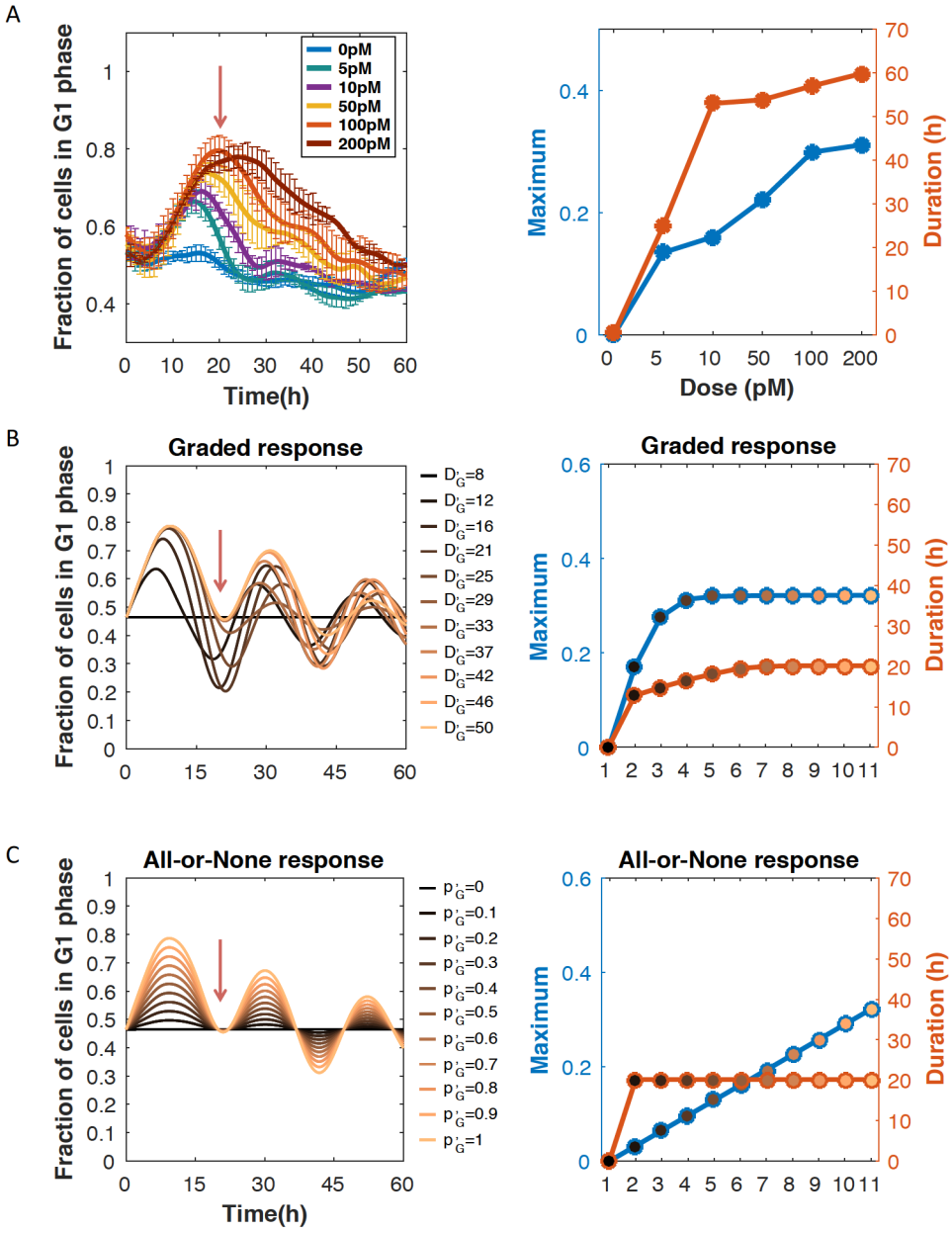


Figure 3.7 Comparison between the model simulations and experimental data. (A) The experimental data (The same data as shown in Figure 2.8). (B) Model simulation results of graded response. Left: the fractions of cell in G1 phase over time (The same as Figure 3.6A); (C) Model simulation results of all-or-none response. Left: the fractions of cell in G1 phase over time (The same as Figure 3.6B). The red arrow indicates the time point at 20h. (B) and (C) Right: the maximums and the durations of the curves in the left panel. The circles (right panel) correspond to the curves with the same color (left panel).

The Model 1 and Model 2 were simulated based on the assumption that cells in S/G2/M phase are unaffected by the cytostatic effect of TGFβ. The inconsistence between the model predictions and experimental data (Figure 3.7 red arrow) indicated that this assumption was incorrect and cells in S/G2/M phase are affected.

3.2.2.2 Model simulation supports the all-or-none response mode

Next, we simulated Model 3 and Model 4, based on the assumption that the cells in S/G2/M phase are affected by the cytostatic effect of TGFβ. The cells in G1 phase are arrested immediately (with average arrested G1 phase duration D'_{1stG}) and the cells in S/G2/M phase will be arrested in the G1 phase of the next cell cycle (with average arrested G1 phase duration D'_{2ndG}) (Figure 3.4B).

Taken sensitivity of cells in G1 (G_{Sen}) or S/G2/M phase (M_{Sen}) into consideration, we did three pairs of simulations for comparison (Table 3-9, Figure 3.8, values of parameters and initial G1 phase durations see Appendix Chapter 9:9.2).

Table 3-9 Simulations of Model 3 and Model 4

	Sensitivity	Model 3 Graded response	Model 4 All-or-none response
1	$G_{Sen}= M_{Sen}$ (Figure 3.8 A and D)	$D'_{1stG} = D'_{2ndG}$ The arrested G1 durations of the cells stimulated in G1 or S/G2/M phase were fitting to the same distribution	$p'_G = p'_M$ The cells stimulated in G1 or S/G2/M phase had the same probabilities to be arrested
2	$G_{Sen}< M_{Sen}$ (Figure 3.8 B and E)	$D'_{1stG} < D'_{2ndG}$ The average arrested G1 duration of the cells stimulated in G1 was shorter than cells stimulated at S/G2/M phase	$p'_G < p'_M$ The cells stimulated in G1 phase had the smaller probability to be arrested than in S/G2/M phase
3	$G_{Sen}> M_{Sen}$ (Figure 3.8 C and F)	$D'_{1stG} > D'_{2ndG}$ The average arrested G1 duration of the cells stimulated in G1 was longer than cells stimulated at S/G2/M phase	$p'_G > p'_M$ The cells stimulated in G1 phase had the larger probability to be arrested than in S/G2/M phase

In the three pairs of simulation results of Model 3 and Model 4, we observed that the duration could reach 50-60h in certain conditions of both response modes (Figure 3.8), which represented a characteristic of experimental data. This supported the hypothesis that an inheritable G1 arrest induced by the stimulation during S/G2/M phase.

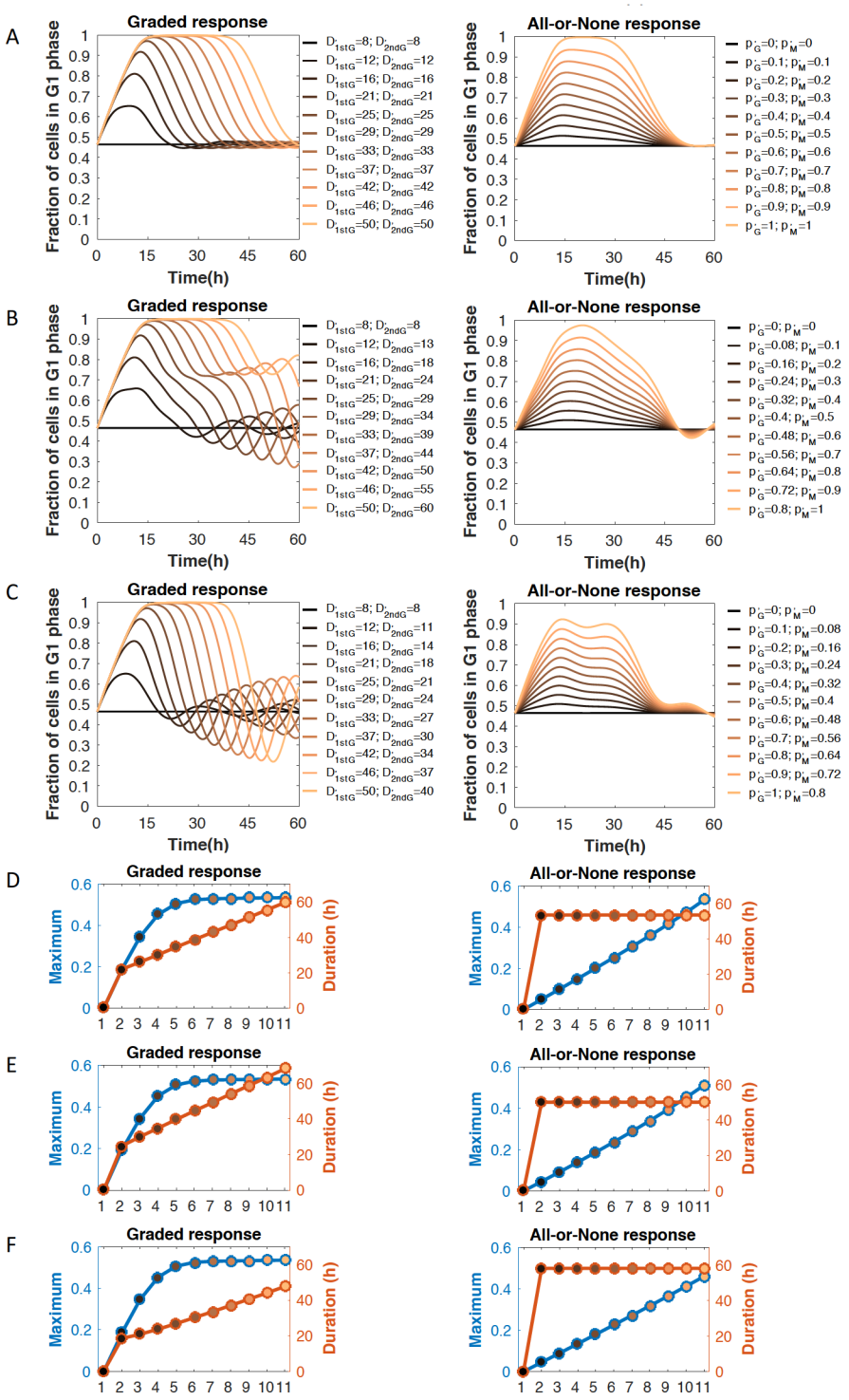


Figure 3.8 Comparison between simulation results of Model3 (graded response) and Model 4 (all-or-none response). (A), (B) and (C) Left: Simulation results of Model 3 (graded response); D'_{1stG}: the mean of arrested G1 phase duration for the cells arrested at 1st G1; D'_{2ndG}: the mean of arrested G1 phase duration for the cells arrested at 2nd G1. Right: Simulation results of Model 4 (all-or-none response); p'_G : the probability of being arrested when stimulated in G1 phase; p'_M: the probability of being affected when stimulated in S/G2/M phase. (D), (E) and (F) The maximums and durations of the simulation results in (A), (B) and (C), respectively. The circles correspond to the curves with the same color.

Notably, there was a significant difference between graded (Model 3) and all-or-none (Model 4) response modes: the durations of the graded response groups kept increasing upon the increasing concentration of ligand treatment, while they were the same of all-or-none response groups.

Experimental data (Figure 3.7A) show that the duration was shorter of the response to 5pM TGFβ stimulation (low dose), which might due to the variations and noise in the experimental measurements. However, the durations were similar of the responses to 10, 50, 100 and 200pM TGFβ stimulation. The Model 3 could not reproduce the responses with this characteristic.

Taken together, only the model 4 represented the characteristics of experimental data (Table 3-10), supporting the hypotheses that (1) TGFβ induces all-or-none G1 arrest response; (2) The cells in G1 phase are arrested immediately and the cells in S/G2/M phase will be arrested in the G1 phase of the next cell cycle.

Table 3-10 PBM Model 4 represents the characteristics of experimental data

	Graded response	All-or-None response
Cells in S/G2/M are unaffected	Model 1 ⛔	Model 2 ⛔
Cells in S/G2/M are arrested in 2nd G1 phase	Model 3 ⛔	Model 4 ✔

* ⛔: the model can not represent the characteristics of experimental data; ✔ : the model can represent the characteristics of experimental data.

Chapter 4

Single cell analysis of cell-cycle outcomes induced by TGFβ

Through the model simulations, we found some interesting single cell behaviors hidden behind the population average: (1) the inheritable G1 arrest in the next cell cycle induced by TGFβ at S/G2/M phase; (2) TGFβ induces all-or-none response.

In this chapter, we confirmed the model predictions by single cell data from live-cell imaging experiments. In addition, we found that G1 arrest in response to TGFβ stimulation is both phase-dependent and dose-dependent. Interestingly, the lengthened duration does not depend on the timing within each cell cycle phase.

We also explored the correlation between Smad2 dynamics and G1 arrest duration in individual cells, and found the correlation between them is weak.

4.1 Inheritable G1 arrest induced by TGFβ during S/G2/M phase

To reveal the effect of TGFβ on each cell cycle phases at the single cell resolution, we tracked the cells and recorded the cell cycle trajectories of individual cells with 3h TGFβ1 treatment (Doses: 0, 10, 50 and 100pM) (Figure 4.1).

The G1 and S/G2/M durations of the cell cycle upon the stimulation were recorded. As the model predictions predict delayed G1 arrest in the next cell cycle when stimulated at S/G2/M phase, the G1 duration of the following cell cycle (2nd G1 duration) was also recorded.

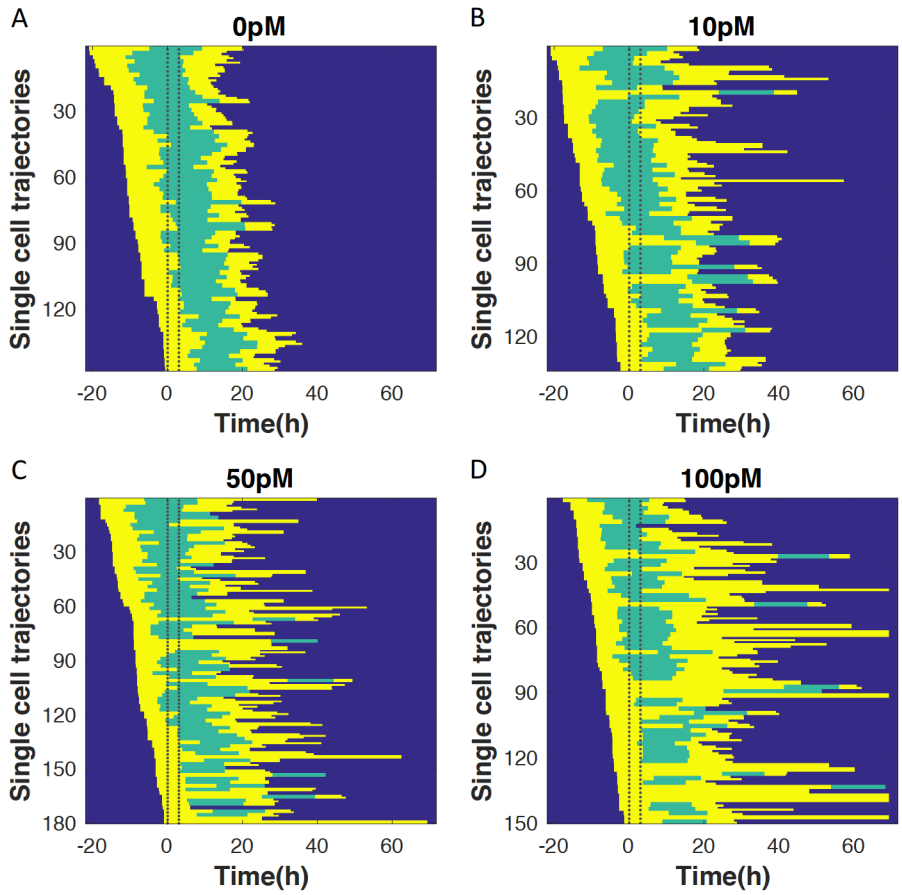


Figure 4.1 Single cell analysis of cell cycle progression. (A)-(D) Asynchronously cycling cells were treated with 0, 10, 50 or 100pM TGFβ1 for 3h. The time of adding TGFβ is set as zero. The duration of TGFβ treatment is shown (Within dash lines). Each row represents the trajectory of an individual cell. Yellow indicates the cell was in G1 phase and green indicates the cell was in S/G2/M phase. Cells were sorted by the time of birth. 148, 134, 180 and 150 cells of the groups in Figure 2.8 were tracked and quantitatively analyzed (order corresponds to 0, 10, 50 and 100pM).

The phase durations of each cell were quantitatively analyzed (Figure 4.2). As the asynchronous proliferating cells were either in G1 or S/G2/M phase upon TGFβ stimulation, we divided the cells into two groups by the phase at which the

cell is treated and compared the phase durations (1st G1, S/G2/M and 2nd G1 phase).

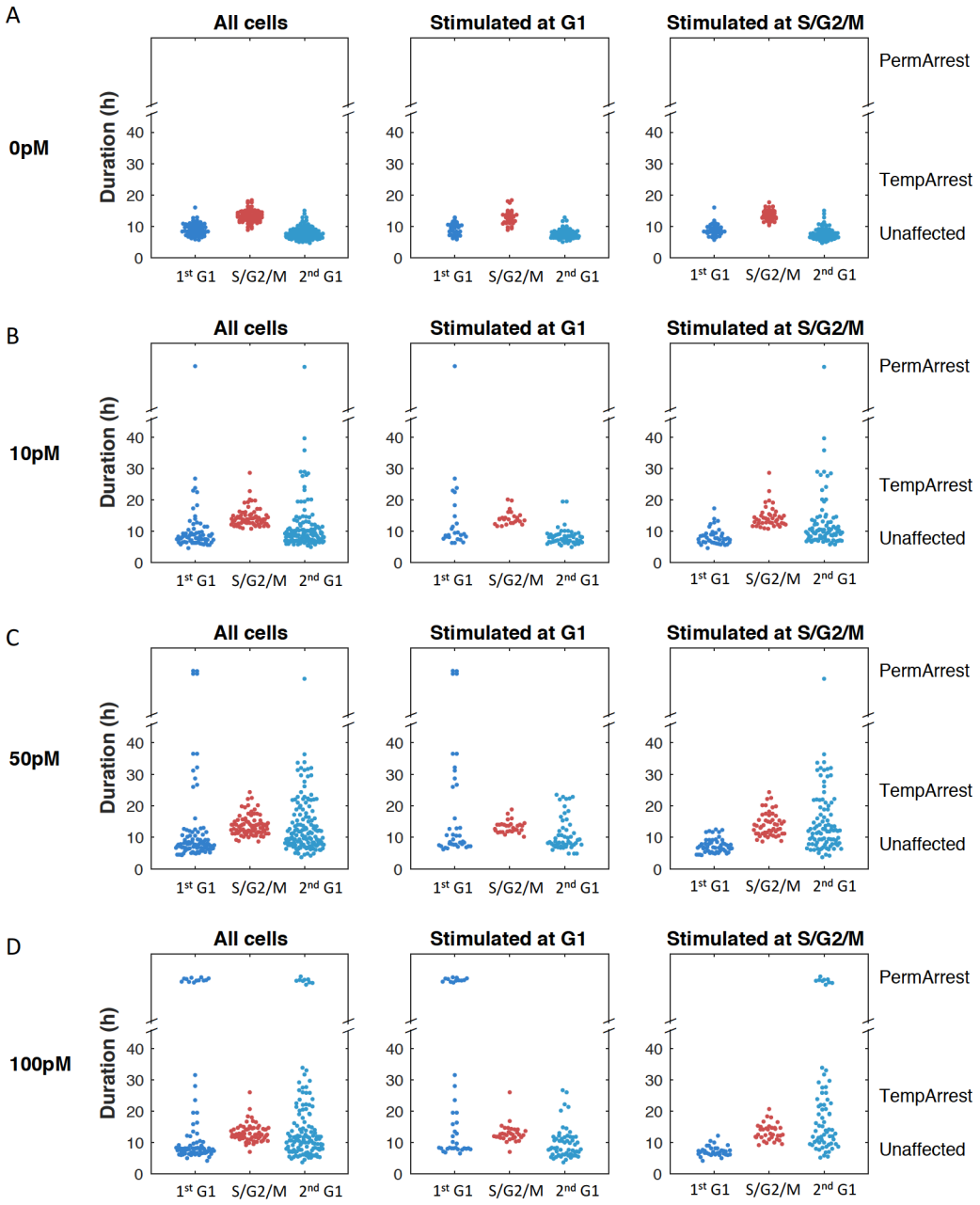


Figure 4.2 Quantitatively analysis of the cell cycle phase durations. (A)-(D) The different cell cycle phase durations of the cells upon 0, 10, 50 or 100pM TGFβ1 treatment in Figure 4.1. The cells were divided into two groups: stimulated at G1 phase (Middle panel) and stimulated at S/G2/M phase (Right panel). The G1 (1st G1) and S/G2/M duration of the cell cycle upon TGFβ1 treatment, and the G1 duration in the following cell cycle (2nd G1) were shown separately.

Results showed that it was only G1 phase that was prolonged upon TGFβ stimulation, which is consistent with the knowledge that TGFβ activates cytostatic gene responses and primarily targets G1 event[9].

When exploring the inheritability of the cytostatic responses, we found that although the S/G2/M phase duration didn't change significantly, the following G1 phase (of the next cell cycle) was lengthened for the cells stimulated at S/G2/M phase, but not for most of the cells stimulated at G1 phase.

There were a small number of cells stimulated at G1 phase with lengthened 2nd G1 phase. They turned out to be the cells that were at the end of G1 phase when TGFβ was added, and they were mainly in S/G2/M phases during the 3h of treatment (Figure 4.3).

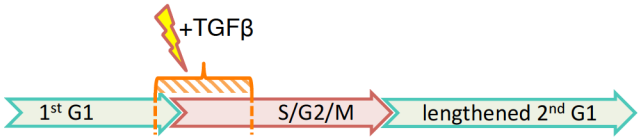


Figure 4.3 Schematic showing the cells with lengthened 2nd G1 phase undergo G1/S transition during TGFβ treatment. The cells are at the late stage of G1 phase when the TGFβ is added and then progress into S/G2/M phase. They are mainly in S/G2/M phases during the TGFβ treatment.

Thus, we draw a conclusion that TGFβ stimulation at G1 phase induces immediate G1 arrest, while stimulation during S/G2/M phase induces inheritable, delayed G1 arrest. In the following studies, we only focused on the 1st G1 duration of the cells stimulated at G1 phase and the 2nd G1 duration of the cells stimulated at S/G2/M phase.

4.2 All-or-none G1 arrest outcomes induced by TGFβ

Next, to distinguish between the graded or all-or-none response, we plotted the G1 phase durations of the cells upon TGFβ treatment in G1 or S/G2/M phase (Figure 4.4).

Without TGFβ stimulation, most of the cells stay in a relatively concentrated group (control group). In contrast, cells upon TGFβ could be categorized into three different groups according to the length of G1 duration: unaffected group, temporary arrested (TempArrest) and permanent arrested (PermArrest) group (Figure 4.4A).

1) Unaffected group: cells with G1 durations (<12.24h) that are similar to that of the control group (within 95% percentile of the G1 durations of the control group).
2) TempArrest group: cells with significantly prolonged G1 phase (12.24-40h).

3) PermArrest group: cells with G1 duration over 40h. For the cells that show prolonged G1 phase, most of them that progress into S phase during the imaging experiments have prolonged G1 phase durations below 40h. However, some of them stay at G1 phase for over 50h, and they are still in G1 phase at the end of imaging experiments. Therefore, there are clearly two categories of arrested cells, and we assign the cells with G1 phase durations longer than 40h to the PermArrest group.

Calculating the percentage of cells in each group upon various doses stimulations, we found that the fraction of affected cells (including TempArrest and PermArrest) increases when raising the TGFβ dose (Figure 4.4B). Notably, the fraction of temporary arrest cells did not increase any more over 50pM. Instead, more cells shifted to the permanent arrest group. Interestingly, there were still around 40% cells escaping from the growth-inhibitory function of TGFβ at the concentration of 100pM, which is a saturated dose and supposed to strongly inhibit cell cycle progression (Figure 2.8). The result of three different cell-cycle arrest outcomes suggested that TGFβ triggered all-or-none cell cycle arrest responses.

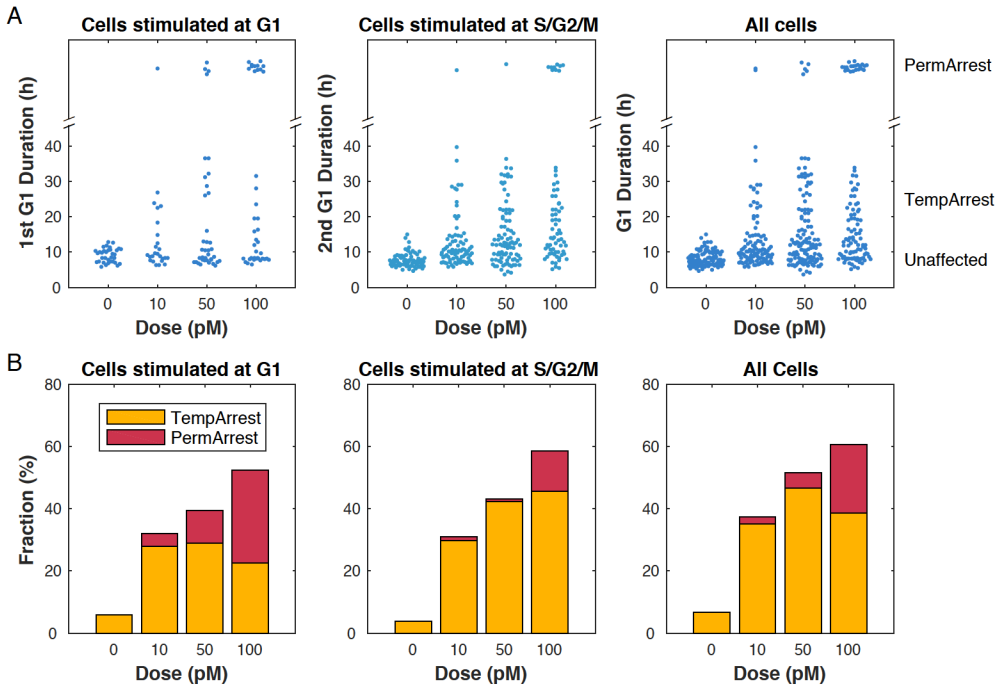


Figure 4.4 TGFβ triggers all-or-none cell cycle arrest response. (A) There are three outcomes depending on the G1 duration: unaffected, temporary (TempArrest) and permanent arrest (PermArrest). Left: the 1st G1 duration of the cells stimulated at G1 phase; Middle: the 2nd G1 duration of the cells stimulated at S/G2/M phase; Right: plot the data in Left and Middle panels together. 1st G1: the G1 phase in the current cell cycle upon TGFβ treatment; 2nd G1: the G1 phase in the next cell cycle following the cell cycle upon TGFβ treatment. (B) The fractions of temporary/permanent arrested cells in response to 0, 10, 50 or 100pM TGFβ1 stimulation in (A).

4.3 TGFβ-induced G1 arrest is phase-dependent and dose-dependent

TGFβ could induce G1 arrest no matter the cell is at G1 or S/G2/M phase. We asked if cells in G1 and S/G2/M phases are equally sensitive to TGFβ stimulation.

The fractions of affected cells (Including TempArrest and PermArrest) are similar for stimulations at G1 phase and S/G2/M phase (Figure 4.4B). However, compared with stimulations at S/G2/M phase, stimulations at G1 phase always induce higher fraction of permanent arrest cells for the same TGFβ concentration (Figure 4.5). This result suggests that the cells in S/G2/M phase are less sensitive to TGFβ stimulation.

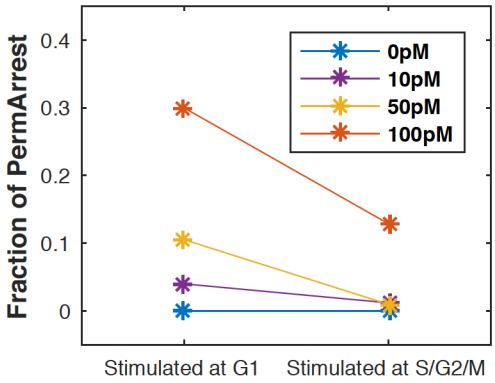


Figure 4.5 The sensitivity to the concentration of TGFβ is cell cycle phase dependent. The fraction of PermArrest cells for stimulations at G1 or S/G2/M phase by 0, 10, 50 or 100pM TGFβ1.

When zooming into each cell cycle phase, the asynchronous cells were at different stages of the phase when stimulation happened (Figure 4.1). Therefore, we would like to explore whether the cells in certain stages are more sensitive to the TGFβ. Here we used 'the time spent before stimulation within the phase' to indicate the stage upon stimulation within the cell cycle phase (Figure 4.6).

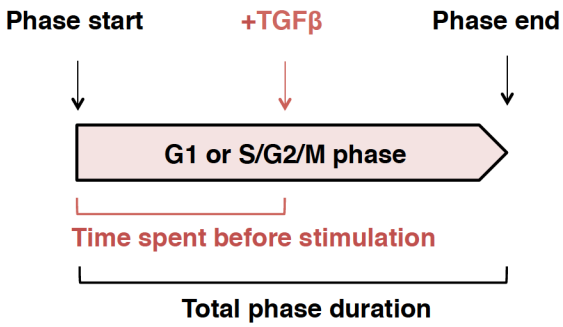


Figure 4.6 Schematic definitions of the 'time spent before stimulation' and the 'total phase duration'. For each cell stimulated in G1 phase or S/G2/M phase, the total phase duration means the period from the onset to the end of the phase; the time spent before stimulation means the period from the onset of the phase to the time point of adding TGFβ.

Upon low dose of TGFβ1 stimulation (10pM), we found that few cells were affected in the first 5h in S/G2/M phase (Figure 4.7B). Though there were several cells categorized as TempArrest, their G1 phase durations were quite close to the up-bound of the G1 phase duration of the unaffected group. This could be caused by the statistical threshold setting (95 percentile of the control duration) so that a small amount of unaffected cells could be categorized into TempArrest Group.

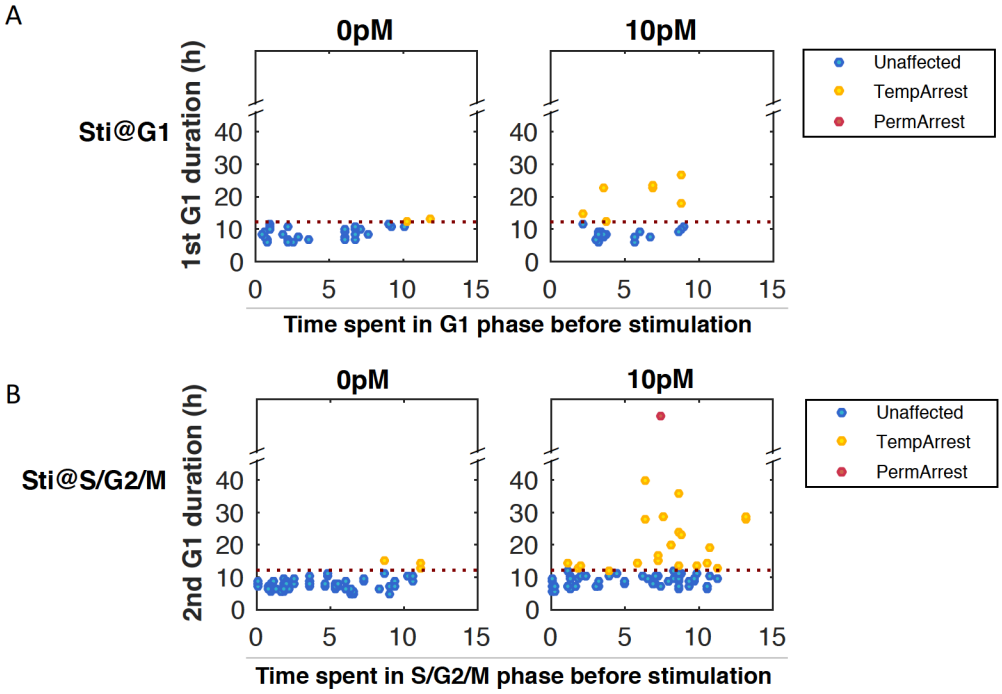


Figure 4.7 The G1 duration and the timing of TGFβ stimulation within the phase. The G1 arrest outcomes are plotted against the timing of 0/10pM TGFβ doses stimulation within the G1 (A) or S/G2/M phase (B). The red dash line: 95 percentile of control duration.

To confirm the existence of a insensitive time window, we analyzed more cells stimulated in the first 5h of S/G2/M phase by varied doses of TGFβ (0, 10, 20, 30, 50 and 100pM). We found that upon low doses of TGFβ stimulation (10/20pM), few cells were affected in the first 5h of S/G2/M phase. However, the G1 arrest response was triggered above the dose of 30pM (Figure 4.8).

Although we did not have specific indicator for S phase, based on the knowledge in literature, this 5h insensitive time window is most likely to be the S phase by considering the average S phase duration (~5h) of HaCaT cells [139]. This switch-like response suggests that the sensitivity of S phase to TGFβ-induced G1 arrest is dose dependent. In contrast to the insensitivity of S phase to low doses of TGFβ, cells in G1 or G2/M phase were arrested upon 10pM TGFβ treatment (Figure 4.7). This confirmed that the sensitivity to the TGFβ is cell cycle phase dependent.

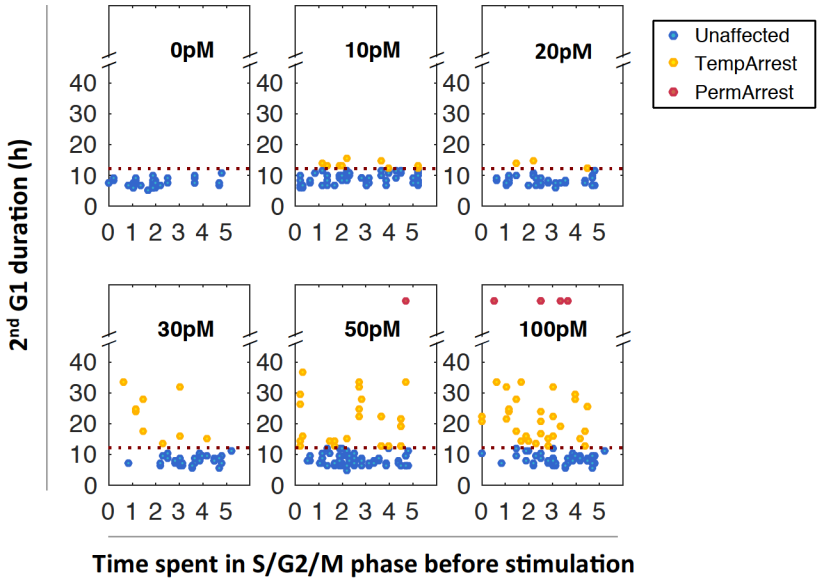


Figure 4.8 Switch-like response in the first 5h in S/G2/M phase. For cells stimulated in the first 5h of S/G2/M phase, cell-cycle arrest outcomes in response to various doses of TGFβ. Ultrasensitivity (switch-like arrest response) was observed: G1 arrest response was triggered above the dose of 30pM. The red dash line: 95 percentile of control duration.

Taken together, both the timing and dose of TGFβ stimulation are critical for its cytostatic effect: compared with stimulations at S/G2/M phase, stimulations at G1 phase always induce a higher fraction of permanent arrest cells under the same concentration of TGFβ treatment. For stimulations at S phase, low dose of TGFβ does not induce G1 arrest.

4.4 TGFβ-induced G1 arrest and the timing of stimulation within the phase

For high doses of TGFβ treatment, G1 arrest was induced upon stimulation in G1, S and G2/M phase. We asked if the G1 arrest outcome depends on the timing of TGFβ treatment within a cell cycle phase. The cells upon 50pM TGFβ treatment were analyzed quantitatively (Figure 4.9).

Firstly, for the temporarily arrested cells, we explored the correlation between their G1 durations and the timings upon TGFβ stimulation within each phase (In terms of the Pearson's rho). The correlations were low (Figure 4.9B).

Next, we examined the fractions of arrested cells (either temporary or permanent arrested) upon TGFβ stimulation on early, middle or late stage within each phase. By fitting the cell phase durations to lognormal distributions, the theoretical distributions were calculated. Each phase was roughly divided into

three equal partitions (3 stages: early, middle or late) according to the theoretical distribution. The cells were assigned to each stage within the phase based on the time spent in that phase upon stimulation.

In G1 phase, cells in the early stage are less likely to be affected by TGFβ than in middle and late stages. In S/G2/M phase, the affected fractions are similar among early, middle and late stages (Figure 4.9C).

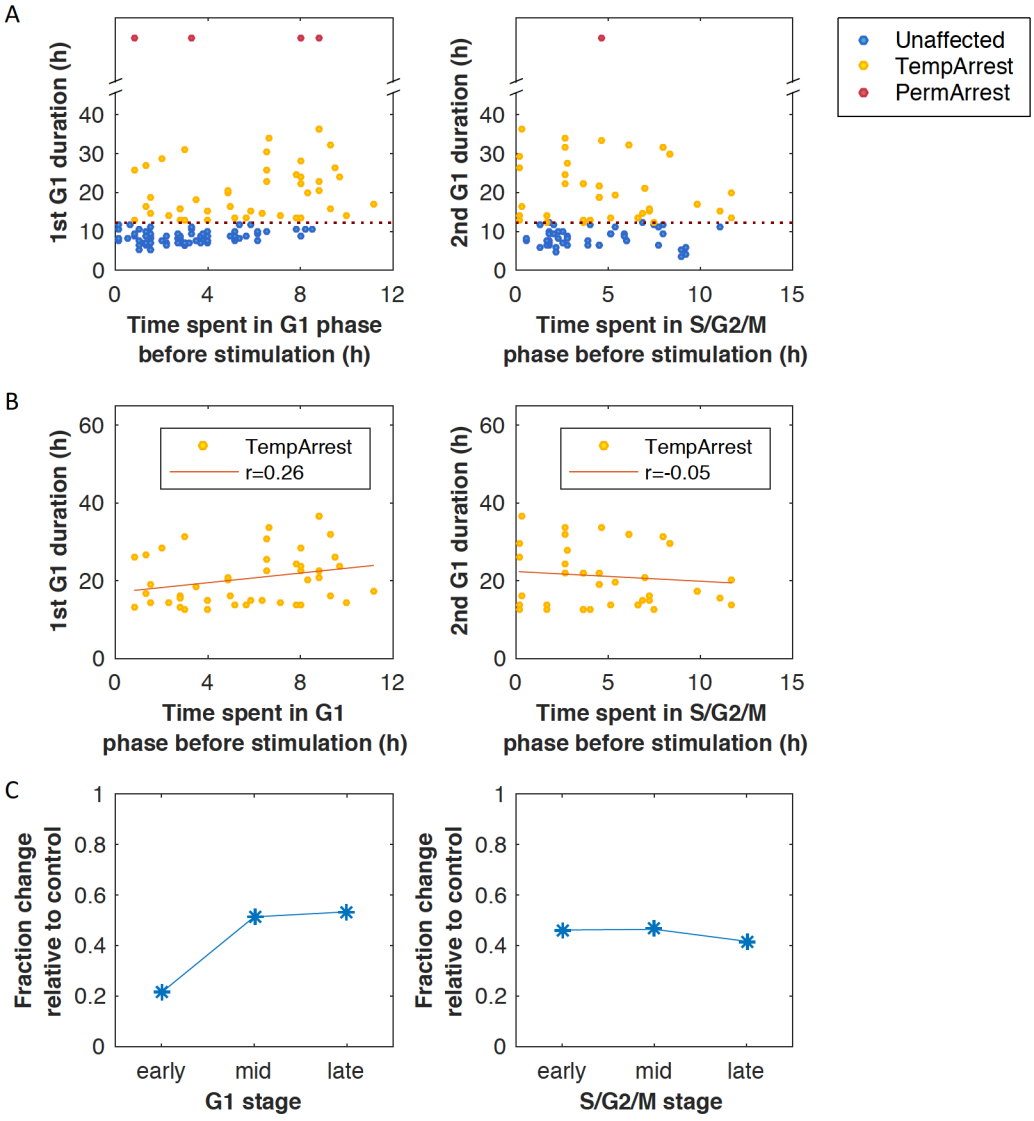


Figure 4.9 The G1 arrest and the timing of TGFβ stimulation within each phase. (A) The cell-cycle arrest outcomes (left: 1st G1 duration; right: 2nd G1 duration) relative to the timing of 50pM TGFβ1 stimulation for cells stimulated at G1 (left) or S/G2/M phase (right). The red dash line: 95 percentile of control duration. (B) The 1st or 2nd G1 duration of temporary arrested cells as a function of the time spent in G1 or S/G2/M phase, respectively. Red line represents the linear regression of the data. r, Pearson's correlation coefficients. (C) The cell-cycle arrest outcomes when stimulated at the early, middle or late stage within the G1 or S/G2/M phase. The fraction change relative to control: the difference between the fractions of arrested cells (Temporary arrested and Permanent arrested) upon 0pM and 50pM treatment.

In summary, when stimulated at high doses of TGFβ, the G1 durations of temporary arrest cells are independent of the timing upon stimulation within the cell cycle phase. However, the fraction of affected cells is dependent of the stage within G1 phase upon stimulation: cells in the early stage of G1 phase are less likely to be affected by TGFβ.

4.5 Correlation between Smad2 nuclear accumulation and G1 arrest duration

As the Smad signaling plays a central role in the canonical TGFβ signaling, we suspected that the TGFβ-induced cell cycle arrest is mediated through the Smad2 signaling. Besides imaging the cell cycle progression in response to various doses of TGFβ1, we simultaneously recorded the nuclear accumulation of EYFP-Smad2 in single cells.

At the average level, the nuclear EYFP-Smad2 signal increased after stimulation and reached the maximum in 2h. At the population level (Figure 4.10, thick line), below the saturation dose of ~50pM, higher dose could induce higher level of Smad2 nuclear translocation. Notably, at the single cell level, obvious heterogeneity exists in EYFP-Smad2 dynamics (Figure 4.10, thin lines).

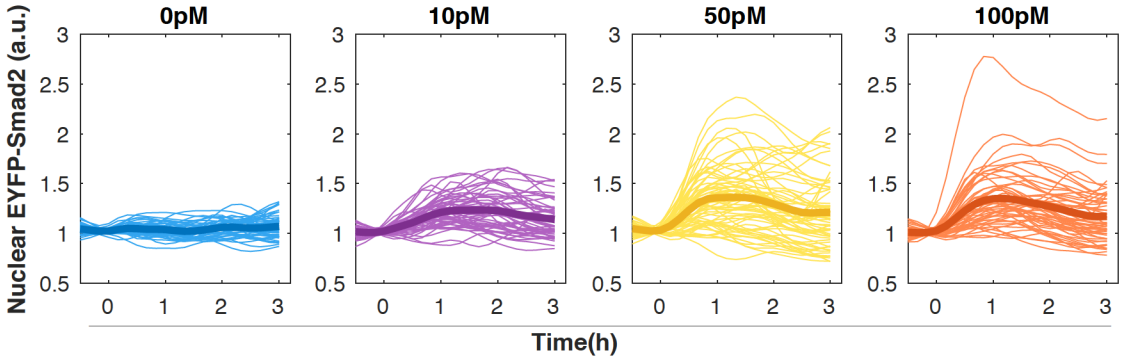


Figure 4.10 Time-resolved analysis of nuclear EYFP-Smad2 upon 3h of various doses of TGFβ stimulation. The time of adding TGFβ was set as zero. The EYFP-Smad2 dynamics of individual cells (thin lines) and population average (thick line) were shown.

Next, we would like to study the correlation between the Smad2 signaling dynamics and the G1 phase prolongation.

Specifically, for quantifying the dynamics of Smad2 nuclear accumulation, we took into consideration the amplitude (the maximum level/basal level) and the activation time (the duration from the time of adding TGFβ until the nuclear EYFP-Smad2 reaching the maximum) (Figure 4.11).

For quantifying cell cycle arrest, we only took into account the affected G1 phase duration, i.e. 1st G1 duration of cells stimulated at G1 phase; 2nd G1 duration of the cells stimulate at S/G2/M phase (Figure 4.2).

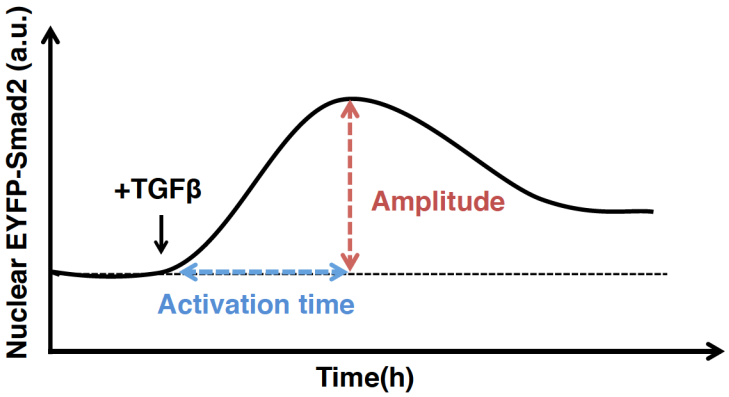


Figure 4.11 Schematic definitions of Amplitude and Activation time of the nuclear EYFP-Smad2 signaling dynamics. The curve describes the dynamics of nuclear EYFP-Smad2 in a cell. The amplitude: the fold change of the maximum compared with the basal level. The activation time: the duration from the time of TGFβ stimulation until the signal reaching the maximum.

There was no obvious correlation between the Smad2 nuclear accumulation and the time spent in cell cycle before stimulation (Figure 4.12), which confirmed that the Smad2 dynamics in response to TGFβ1 is not determined by the cell cycle stage[90].

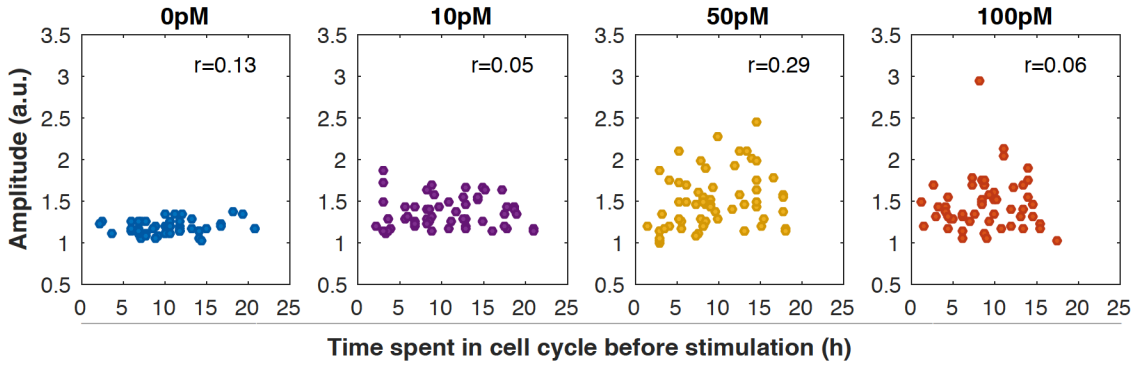


Figure 4.12 Scatter plots of correlations between the EYFP-Smad2 signaling amplitude and the timing of TGFβ stimulation in the cell cycle. r, Spearman's correlation coefficients.

To explore if the G1 duration is determined by the Smad2 signaling dynamics, we analyzed the correlation between the G1 duration and the Smad2 signaling dynamics (Amplitude or activation time). However, we did not observe obvious correlation between them (Figure 4.13).

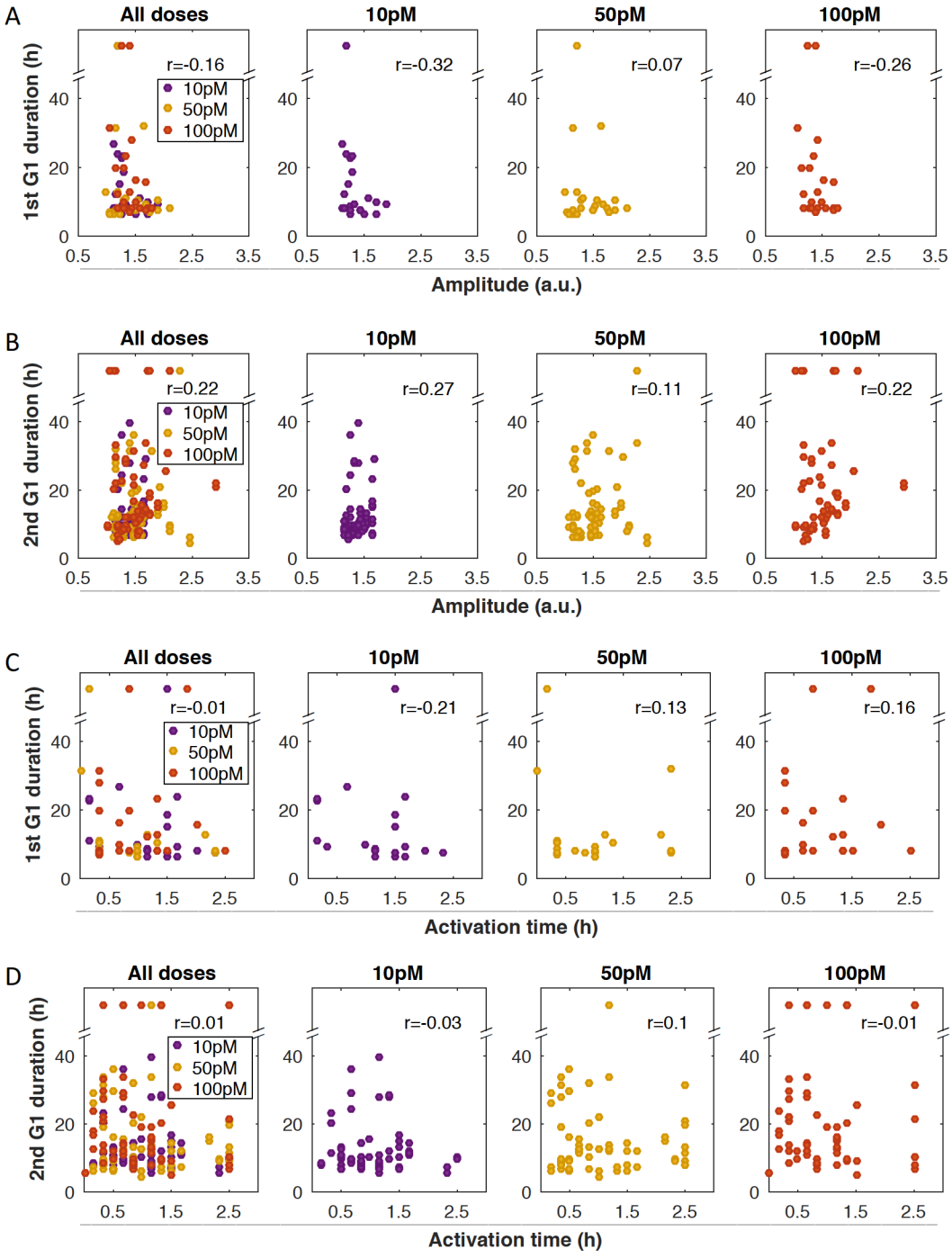


Figure 4.13 Weak correlation between Smad2 nuclear accumulation and G1 arrest duration. (A) Scatter plots showing the correlation between the EYFP-Smad2 dynamics amplitude and the 1st G1 duration upon various stimulus levels for cells treated at G1 phase. (B) Scatter plots showing the correlation between the EYFP-Smad2 dynamics amplitude and the 2nd G1 duration upon various stimulus levels for cells treated at S/G2/M phase. (C) Scatter plots of correlations between Activation time and the 1st G1 duration upon various stimulus levels for cells treated at G1 phase. (D) Scatter plots of correlations between Activation time and the 2nd G1 duration upon various stimulus levels for cells treated at S/G2/M phase. The cells were treated upon varied concentration of TGFβ1 as indicated. r, Spearman's correlation coefficients.

4.6 The levels of p21 in individual cells after TGFβ treatment

In epithelial cells, one of the most important TGFβ-induced cytostatic effects is the activation of p21[9], which is one of the CDK inhibitors. We quantitatively analyzed the p21 level in cells after 3h of TGFβ treatments. The p21 levels of individual cells were varied and the average p21 levels of the cells increased after TGFβ treatments (Figure 4.14).

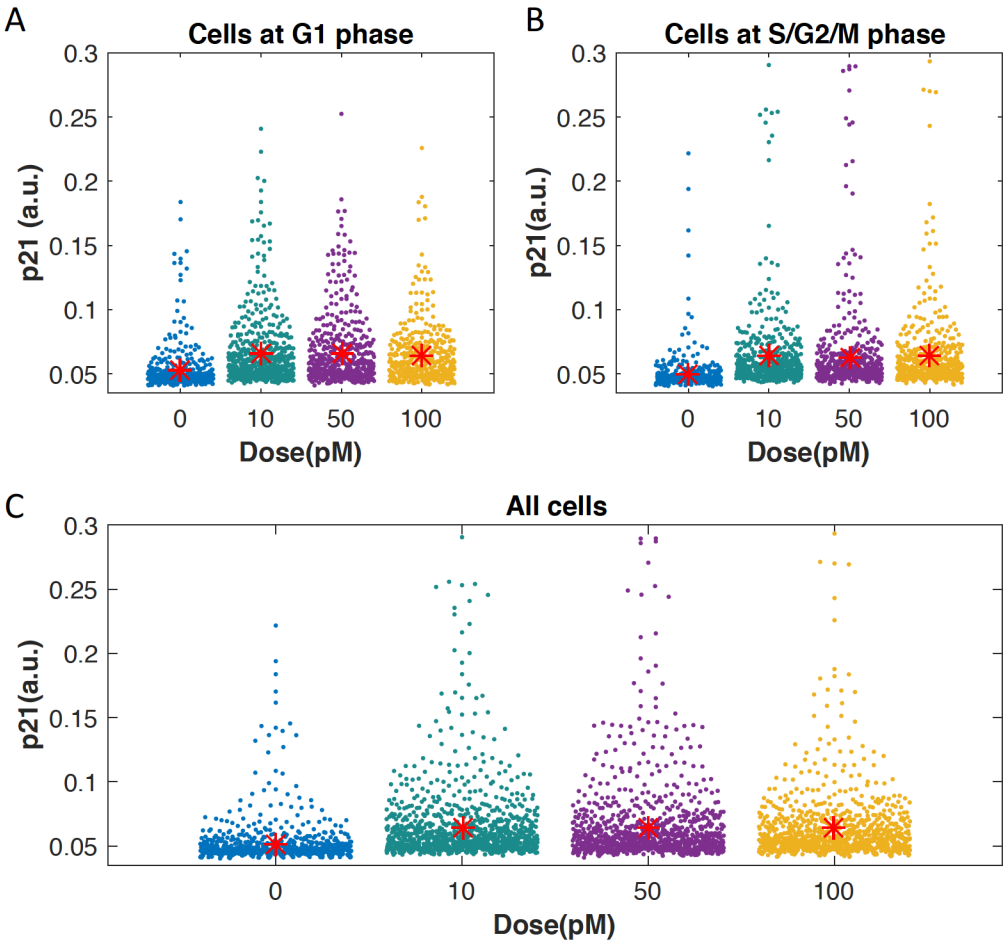


Figure 4.14 TGFβ induces heterogeneous elevated level of p21. The cells were treated by 3h TGFβ (0, 10, 50 or 100pM) and fixed. Then the cells were immunostained for p21 and imaged. The fluorescent signal of p21 was quantitatively analyzed for each cell. The cell cycle phase was indicated by mCherry-Geminin(1/110) of each cell. (A) Cells were in G1 phase; (B) Cells were in S/G2/M phase; (C) plot the data in (A) and (B) together. 877, 1068, 1017 and 924 cells were analyzed (order corresponds to 0, 10, 50 and 100pM). Red asterisks: mean of the p21 signals.

Chapter 5

Modeling of single cell behaviors upon multiple TGFβ stimulations

Combining computational model simulations with experimental analysis, we found that cell cycle arrest in response to single pulse TGFβ stimulation is all-or-none response. In this chapter, we ask the question: how does cells respond to multiple pulses of TGFβ stimulations?

We firstly introduced a Cellular Automaton Model (CAM) to simulate growing single cells. Then we applied the CAM to test different assumptions. Strikingly, we found that the cells are in a refractory state after a pulse of TGFβ stimulation, in which the cells do not response to TGFβ anymore. Based on this knowledge, we proposed a hypothesis to explain the strategy that the cells use to escape the cytostatic effect of TGFβ stimulations.

5.1 Cellular Automaton Model (CAM) for single cells

We could use PBM to simulate cells response upon multiple pulses of TGFβ stimulations by modifying the PBM. There would be five different groups of cells generated upon one pulse of TGFβ stimulation: (1) unaffected, stimulated at G1 phase and (2) temporary arrested or (3) permanent arrested, stimulated at S/G2/M phase and (4) temporary arrested or (5) permanent arrested (Figure 5.1).

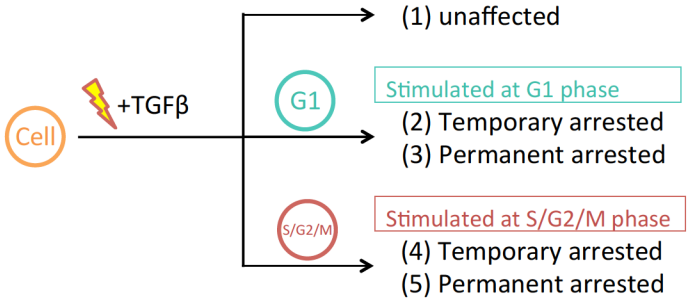


Figure 5.1 Five groups of cells after one pulse of TGFβ stimulation. Cells could be divided to five groups based on their fates and cell cycle phase at the time of TGFβ stimulation.

Thus for n pulses of stimulations, we would need to divide the cells into 5^n groups. As one set of ODEs was used to describe one group of cells in PBM, we should use 5^n sets of ODEs in the model, which is too complicate, e.g. for 3 pulses of stimulation, we would have 125 sets of ODEs.

Here we employed another model, Cellular Automaton Model (CAM), to simulate the individual cell cycle trajectories and mimic cells upon multiple pulses of ligand treatments[133,140]. Although it is computational expensive to simulate single cell cycle trajectories, it is easy to track the response in each cell upon every treatment.

5.1.1 Model assumptions

The CAM was developed based on the following assumptions:

(1) We divide cell cycle duration into two parts: G1 phase duration and S/G2/M phase duration. Each phase duration (G1 or S/G2/M phase) is assumed to follow a lognormal distribution[137,138].

(2) One cell divides into two cells after the completion of the S/G2/M phase. The duration of G1 or S/G2/M phase are reinitialized for these two newborn cells.

(3) At each time step during the cell cycle, the cell could die with a probability p_{dead} and exit the cell cycle.

5.1.2 Model description

In the CAM, there is a grid of cells, each in one of a number of states. The state of one cell $C_i(t)$ (the i -th cell at time-point t) is characterized by four features: (1) dead or the current cell cycle phase θ; (2) the remaining time of staying at the current phase T_R; (3) the duration of its G1 phase T_G; (3) the duration of its S/G2/M phase T_M (Figure 5.2A).

Each cell is assigned a value for its G1 phase duration and another value for its S/G2/M phase duration[133]. The cell cycle progresses independently. Each cell cycle phase is characterized by a sequence of discrete steps in the model[133,140]. When advancing t by 1, a new generation of all the cells is created according to the fixed rules in the decision tree $D\ (\theta, T_R, T_G, T_M)$ (Figure 5.3). For instance, for one cell at the current state of $C_i(t)(\theta = 1, T_R = 80, T_G = 480, T_M = 720)$, when advancing t by 1, it is updated to $C_i(t+1)$ $(\theta = 1, T_R = 79, T_G = 480, T_M = 720)$(Figure 5.2B).

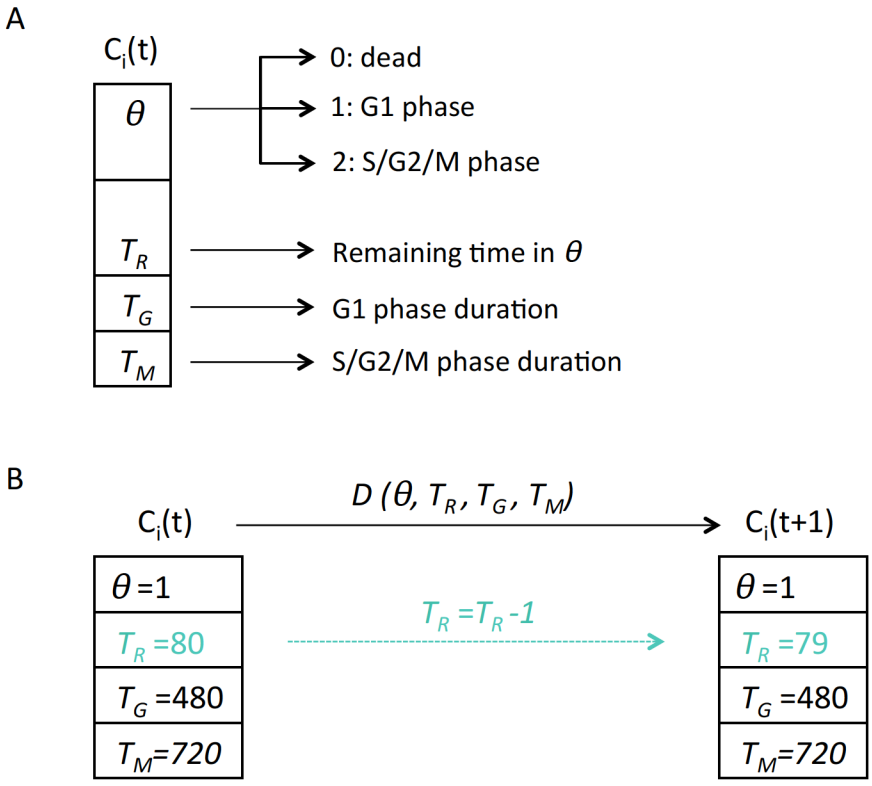


Figure 5.2 The state of cells in the Cellular Automaton Model. (A) The state of one cell $C_i(t)$ is characterized by four features: dead or the current cell cycle phase θ; the remaining time of staying at the current phase T_R; the duration of its G1 and S/G2/M phases T_G and T_M. (B) An example of a cell at state $C_i(t)$ updates to $C_i(t+1)$, according to the rules in decision tree $D\ (\theta, T_R, T_G, T_M)$ in Figure 5.3; the green color: the updated feature.

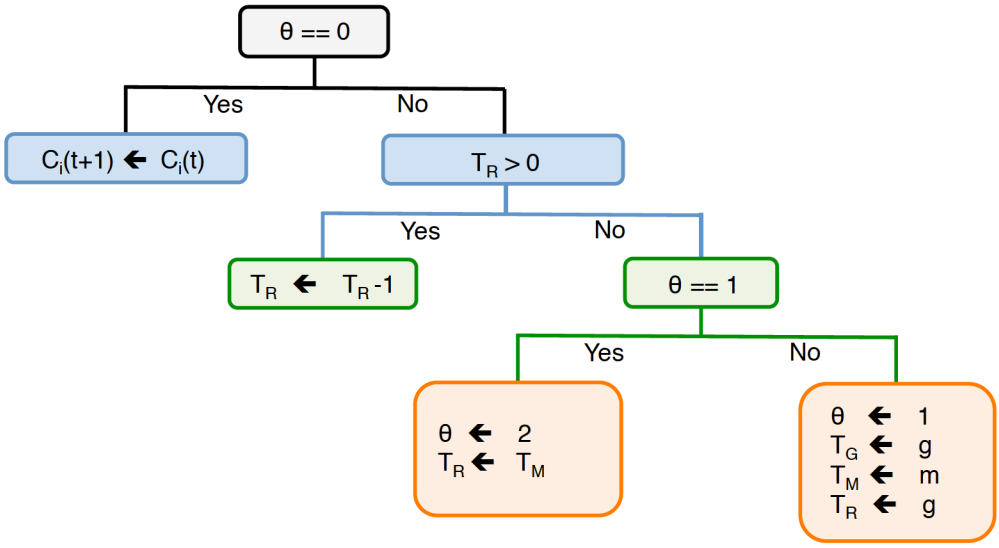


Figure 5.3 Decision tree $D(\theta, T_R, T_G, T_M)$ of the Cellular automaton model. θ ==0,1,2 represent the cell in G1, S/G2/M or death, respectively; $C_i(t)$ is the state of the i -th cell at time-point t; T_R is the remaining time of staying at current phase; T_G and T_M are the duration of G1 phase and S/G2/M phase, respectively; g and m are the newly assigned values for the T_G and T_M, respectively. Black arrow: assign the value to the variable.

In the decision tree, the g and m are newly assigned values for the T_G and T_M of the newborn cell, respectively. Note that an additional cell is produced and assigned with values similarly. In addition, to describe cell death events, we introduced a parameter of p_{dead} into the model. For every generation of cells, we assumed that every cell has the probability (p_{dead}) to die, which is done by assign the θ to 0.

For each cell, the values of T_G and T_M are assigned once it is born. They are sampled from lognormal distributions which were fitted from the experimental data of G1 and S/G2/M phase duration distributions, respectively (Equation 21 and 22).

$$T_{G1} \sim lognomal(\mu_{G1}, \sigma_{G1}) \tag{21}$$

$$T_{S/G2/M} \sim lognomal(\mu_{S/G2/M}, \sigma_{S/G2/M}) \tag{22}$$

Up to now, the system of CAM composed of growing number of cells could progress spontaneously to mimic the cell cycle progression of individual cells in the population.

5.1.3 Steady-state for initial conditions

We assumed that the asynchronous cell population has reached the steady state before TGFβ treatment (the G1 fraction does not change over time). As we

did in PBM, we generated the steady state by setting random values for the initial state of each cell and performing a long-term simulation until the quasi steady state was reached. We stored it as the initial condition for the simulations later on.

5.1.4 The CAM incorporating the effect of TGFβ-induced G1 arrest

Next, we incorporated the effect of TGFβ-induced G1 arrest in CAM. We found that all-or-none G1 arrest response is triggered by TGFβ stimulation and there are five groups of cells generated upon treatment. Thus we employed four parameters to describe the probabilities of cells to be arrested (Arrest Probability parameters, AP parameters): p1, p2, p3 and p4 (Table 5-1).

Table 5-1 Annotation of the four AP parameters

Parameters	Annotation
p1	The probability of TempArrest when stimulated at G1 phase
p2	The probability of PermArrest when stimulated at G1 phase
p3	The probability of TempArrest when stimulated at S/G2/M phase
p4	The probability of PermArrest when stimulated at S/G2/M phase

For a cell in G1 phase, TGFβ stimulation could result in three possible fates (Figure 5.4A): it temporarily arrests in G1 phase and progresses to S phase after a lengthened G1 duration (with a probability of p1); it permanently arrests in G1 phase and does not progress in the cell cycle (with a probability of p2); it is unaffected and progresses to S phase normally (with a probability of 1-p1-p2).

Similarly, for a cell in S/G2/M phase, TGFβ stimulation induces G1 arrest in the next G1 phase after cell division (2nd G1 phase), resulting in three possible fates of the new born daughter cells (Figure 5.4B): it temporarily arrests in G1 phase and progresses to S phase after a lengthened G1 duration (with a probability of p3); it permanently arrests in G1 phase and does not progress in the cell cycle (with a probability of p4); it is unaffected and progresses to S phase normally (with a probability of 1-p3-p4).

Thus, cells upon TGFβ could be categorized into three different groups: unaffected group, temporary arrested (TempArrest) and permanent arrested (PermArrest) group. The effect of TGFβ treatment on the CAM is implemented by applying the four AP parameters. Upon TGFβ stimulation, every cell has certain probabilities of being in TempArrest or PermArrest, depending on their current phase.

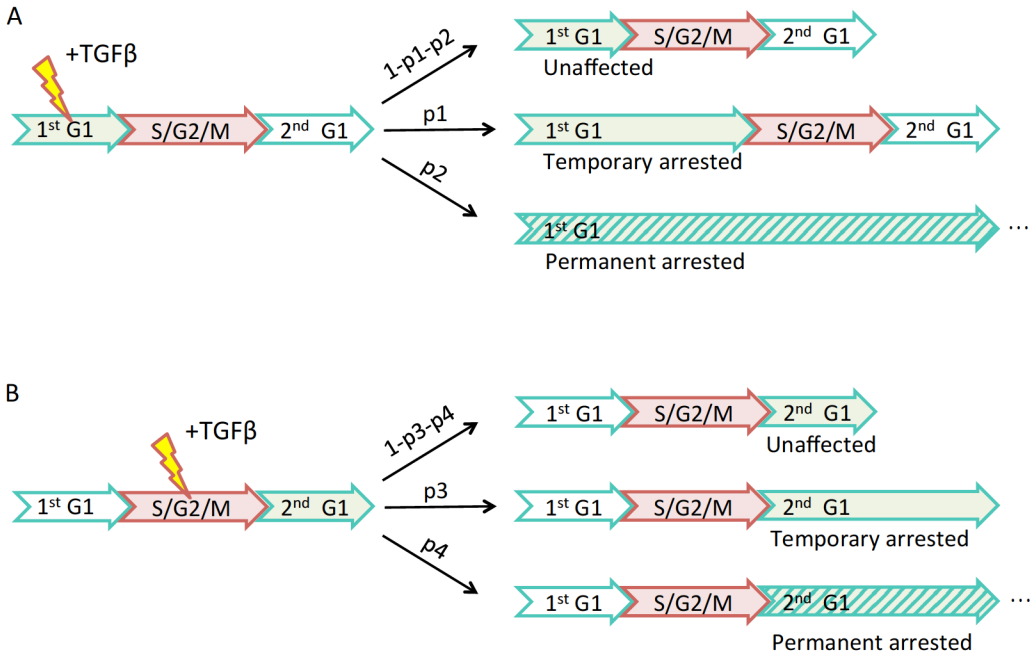


Figure 5.4 Diagram showing the cell fates after TGFβ stimulation. (A) A cell could be temporary arrested (with a probability p1) or permanent arrested (with a probability p2) upon stimulation in G1 phase. (B) A cell could be temporary arrested (with the probability p3) or permanent arrested (with the probability p4) in the next G1 phase after cell division, if the cell is stimulated in S/G2/M phase.

For computational simplification, we introduced an additional label to indicate the state of arrest: the index of arrest I_A into $C_i(t)$, where the value of I_A could be 0, 1 or 2, representing unaffected, TempArrest or PermArrest, respectively (Figure 5.5). Before stimulation, the default value of I_A for all the cells is 0.

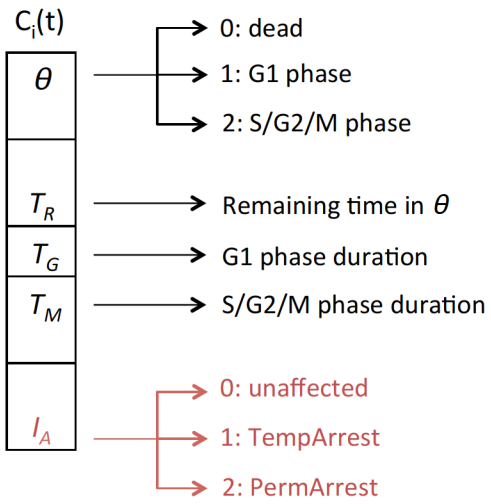


Figure 5.5 The state of cells in the Cellular Automaton Model incorporating the effect of TGFβ-induced G1 arrest. The state of one cell $C_i(t)$ is characterized by five features: dead or the current cell cycle phase θ; the remaining time of staying at the current phase T_R; the duration of its G1 and S/G2/M phases T_G and T_M; state of arrest I_A.

Upon TGFβ stimulation, for cells at G1 phase (i.e. $\theta = 1$), we randomly sample $p1$ and $p2$ of them, and then assign their respective I_A to 1 and 2, respectively. The remaining cells ($(1 - p1 - p2)$ of all the G1 cells) have their default I_A. Upon TGFβ stimulation, for cells at S/G2/M phase (i.e. $\theta = 2$), we randomly sample $p3$ and $p4$ of them, and then assign their respective I_A to 1 and 2, respectively. The remaining cells ($(1 - p3 - p4)$ of all the S/G2/M cells) have their default I_A.

Here we have another modified decision tree $D'(\theta, T_R, T_G, T_M, I_A)$ (Figure 5.6) for incorporating the cytostatic effect of TGFβ stimulation.

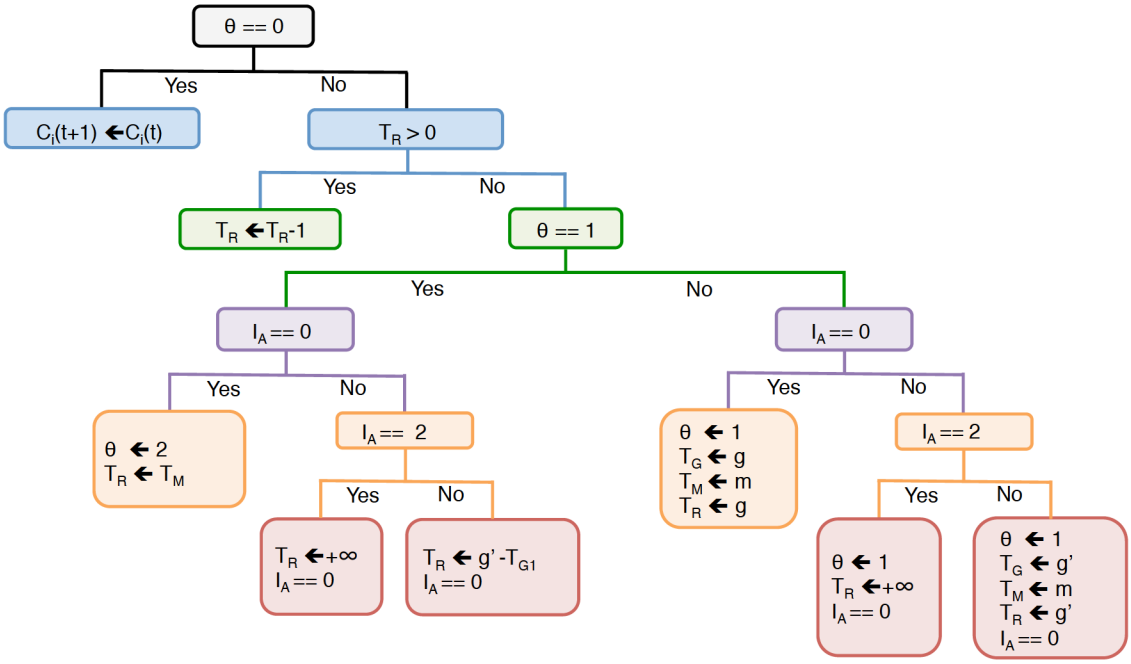


Figure 5.6 Decision tree of the Cellular automaton model (CAM) when incorporating the cytostatic effects. θ ==0,1,2 represent the cell in G1, S/G2/M or death, respectively; $C_i(t)$ is the information of the i -th cell at time-point t; T_R is the remaining time of staying at the current state; T_G and T_M are the duration of G1 phase and S/G2/M phase, respectively; I_A ==0, 1, 2 represent unaffected, TempArrest or PermArrest, respectively; g and m are the newly assigned values for the T_G and T_M, respectively. g′ is the value of arrested G1 phase duration. +∞ is infinity; Black arrow: assign the value to the variable.

In this decision tree, the g and m are newly assigned values for the T_G and T_M, respectively, of the new-born cell; g′ is a value drawn from a lognormal distribution, which is fitted from the experimentally obtained G1 duration distributions of the TempArrest group:

$$T'_G \sim lognomal(\mu'_G, \sigma'_G) \tag{23}$$

By using different sets of AP parameters, the system of CAM is specifically assigned phase durations. These trajectories could then be recorded. In the experiments, we observed an increase of G1 fraction from around 3h after TGFβ treatment. The 3h of delay is also considered in the model simulation.

5.1.5 Estimating the parameters of the CAM

The durations of each cell cycle phase, G1 or S/G2/M, were fitted to experimental data with a lognormal distribution. For each phase, we obtained two parameters: μ and σ, which are the mean and standard deviation of the natural logarithm transformed phase durations.

The parameters in the CAM are shown (Table 5-2).

Table 5-2 Summary of the parameters in CAM

Parameter	Annotation	Value and Unit
μ_G	Parameter μ of the log-normally distributed G1 phase duration	6.1646 min
σ_G	Parameter σ of the log-normally distributed G1 phase duration	0.2218 min
μ_M	Parameter μ of the log-normally distributed S/G2/M phase duration	6.6681 min
σ_M	Parameter σ of the log-normally distributed S/G2/M phase duration	0.1505 min
μ'_G	Parameter μ of the log-normally distributed temporary arrested G1 phase duration	7.3673 min
σ'_G	Parameter σ of the log-normally distributed temporary arrested G1 phase duration	0.2123 min
p1	The probability of TempArrest when stimulated at G1 phase	-
p2	The probability of PermArrest when stimulated at G1 phase	-
p3	The probability of TempArrest when stimulated at S/G2/M phase	-
p4	The probability of PermArrest when stimulated at S/G2/M phase	-

The values of the four AP parameters were experimentally obtained by calculating the percentages of cells in different groups upon stimulation (Figure 4.4). The values of four AP parameters were set according to the dose of ligand (Table 5-3).

Table 5-3 The values of four AP parameters in CAM

Parameters \ Dose of TGFβ(pM)	0	10	50	100
p1	0	0.28	0.2895	0.225
p2	0	0.04	0.1053	0.3
p3	0	0.2976	0.4231	0.4571
p4	0	0.0119	0.0096	0.1286

5.1.6 Model validation

To test if the model could reproduce the data and describe the behaviors of cells upon TGFβ stimulation, we simulated cell responses upon 0, 10, 50 or 100pM treatment (Figure 5.7 and Figure 5.8). Compared with experimental data of single cell analysis (Figure 4.1 and Figure 4.2), the simulation data could reproduce the behaviors of cells upon TGFβ stimulation.

At the population level, we next analyzed the fractions of cells in G1 phase upon TGFβ stimulation over time *in silico*, and then compared the model simulation data with the experimental results. The dynamics were similar for each dose (Figure 5.9).

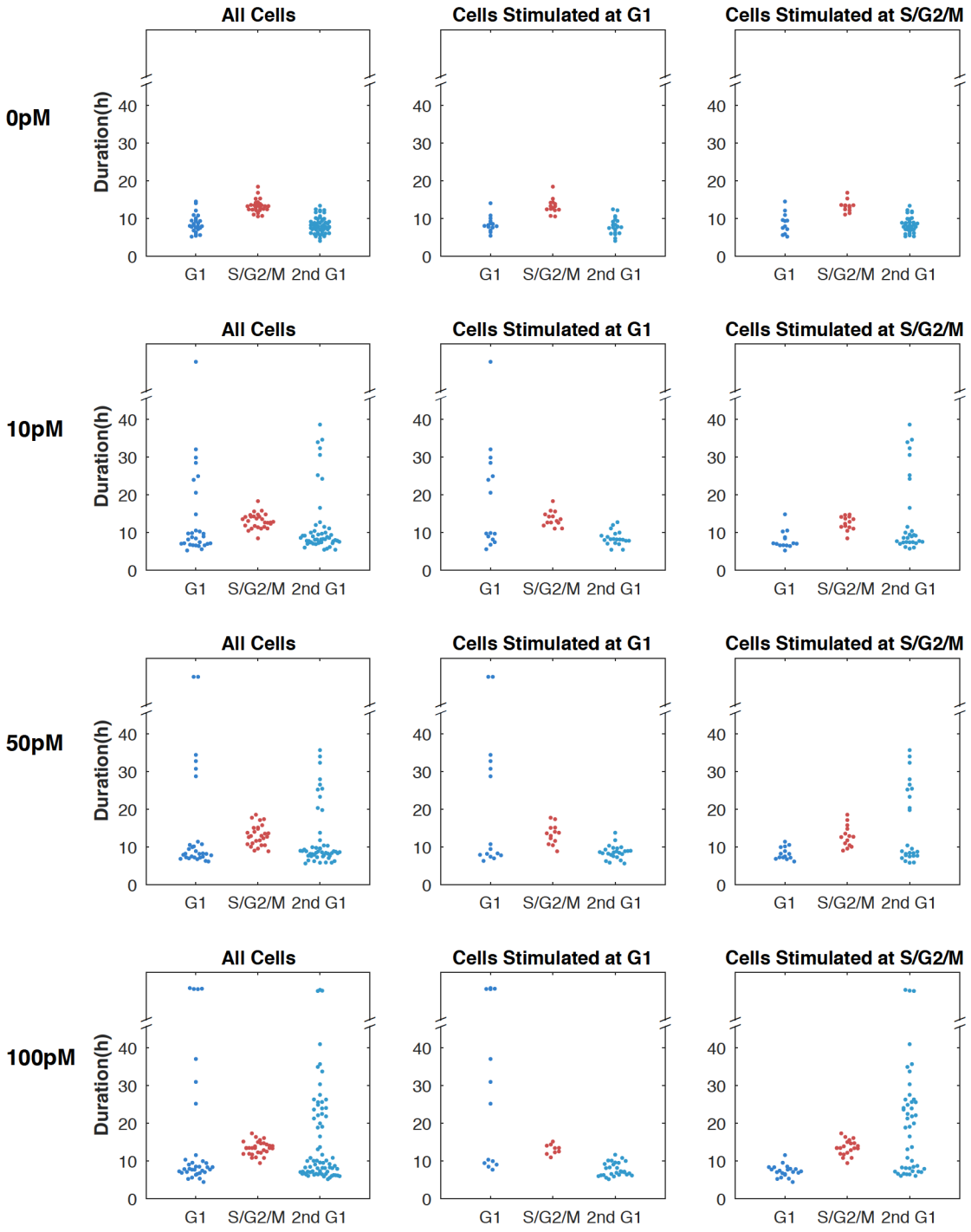


Figure 5.7 Simulations of the cell cycle phase durations. Beeswarm plots of the cell cycle phase durations upon 0, 10, 50 or 100pM TGFβ1 treatment *in silico*. The cells were divided into two groups: stimulated at G1 phase (Middle panel) and stimulated at S/G2/M phase (Right panel). The G1 and S/G2/M duration of the cell cycle upon TGFβ treatment, and the G1 duration in the following cell cycle (2nd G1) were shown separately.

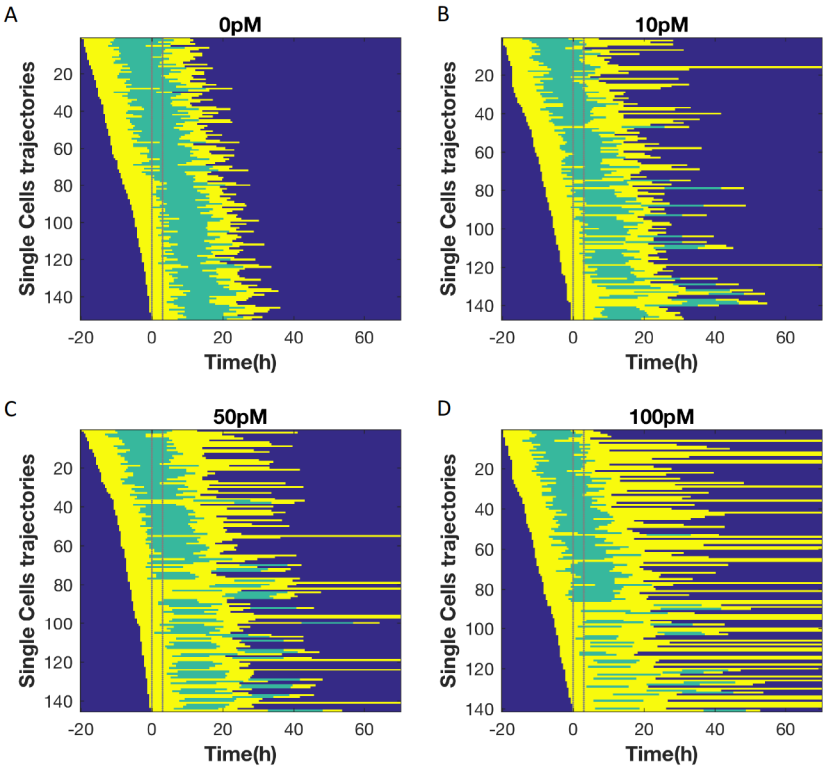


Figure 5.8 Simulations of single cell trajectories. (A)-(D) *In silico*, asynchronously cycling cells were treated with 0, 10, 50 or 100pM TGFβ1 for 3h. The time of adding TGFβ was artificially set as zero. The duration of TGFβ treatment was shown (Within dash lines). Each row represents the trajectory of an individual cell. Yellow color indicates the cell was in G1 phase and the green color indicates the cell was in S/G2/M phase. Cells were sorted by the time of birth.

Notably, in the experiments, we observed that the untreated cells was not always staying at steady state. This might be due to several interferences, e.g., medium change. However, in the model, the untreated cells are always staying at the steady state. To compare model simulation with experimental data, we scaled the control data of the experiments to the steady state of the model and other data (TGFβ treated cells) were scaled accordingly (Details see Chapter 9: 9.3).

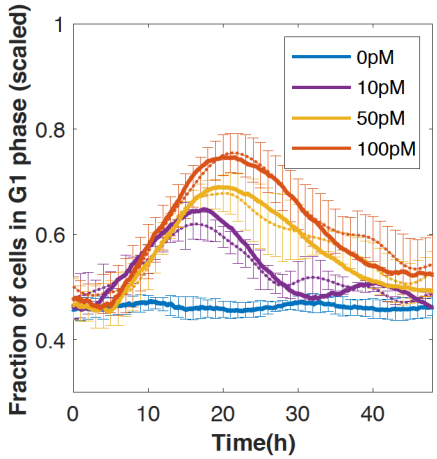


Figure 5.9 Simulation of the dynamics of G1 cells fraction upon 0, 10, 50 or 100pM dose of TGFβ stimulation over time. Asynchronously cycling cells were treated with 0, 10, 50 or 100pM TGFβ1 for 3h. The fractions of cells in G1 phase over time were recorded. Model simulation (solid lines) and experimental data (dash lines) and were shown. The experimental data were scaled accordingly (Details see Chapter 9: 9.3). The experiments were repeated independently at least 3 times. The experimental data were obtained by calculating the average values from 3 to 7 replicates. Error bars: stand error of the mean.

Thus, the CAM could represent the characteristics of cells in experiments at both population and single cell level. We could use CAM to do simulation and prediction.

5.2 Modeling of double TGFβ stimulations

There is always a portion of cells escaping from the TGFβ growth inhibitory regulation (unaffected cells) upon one pulse of stimulation. We would expect that multiple stimulations could arrest more cells than single stimulation. As experimental single cells analysis was time-consuming, we decided to test our hypothesis by simulating single cell cycle trajectories upon TGFβ treatments in CAM and calculate the fraction of cells in G1 phase over time. After that, we could compare the model predictions with experimental data at the population level.

After 1st TGFβ stimulation, there are three states of cells in G1 or S/G2/M phase: unaffected, TempArrest and PermArrest. The assumption of cell cycle arrest induced by 2nd TGFβ stimulation was shown (Figure 5.10 and Figure 5.11).

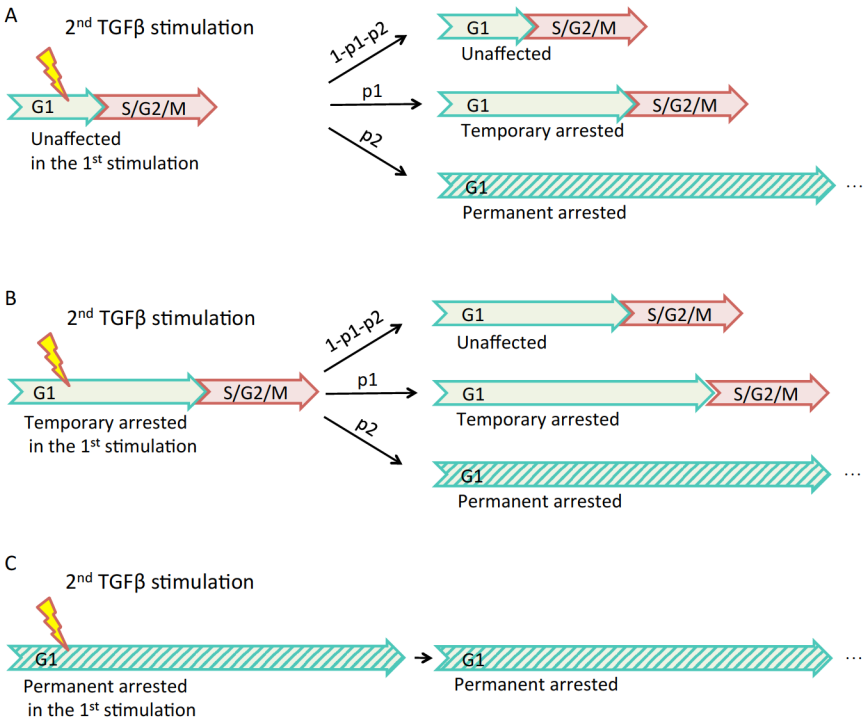


Figure 5.10 Schematic showing the assumption of cell cycle arrest in response to the second TGFβ stimulation in G1 phase in the CAM. There are three groups of G1 cells after the 1st TGFβ stimulation: unaffected, TempArrest and PermArrest. (A) and (B) For unaffected cells and temporary arrested cells, they have the probabilities to be unaffected, temporary arrested or permanent arrested. After the 2nd stimulation, for unaffected cells, they keep the same length of G1 phase as before the 2nd stimulation; for temporary arrested cells, they have a longer G1 phase; for permanent arrested cells, they stay in G1 phase. (C) For the permanent arrested cells, they stay in G1 phase.

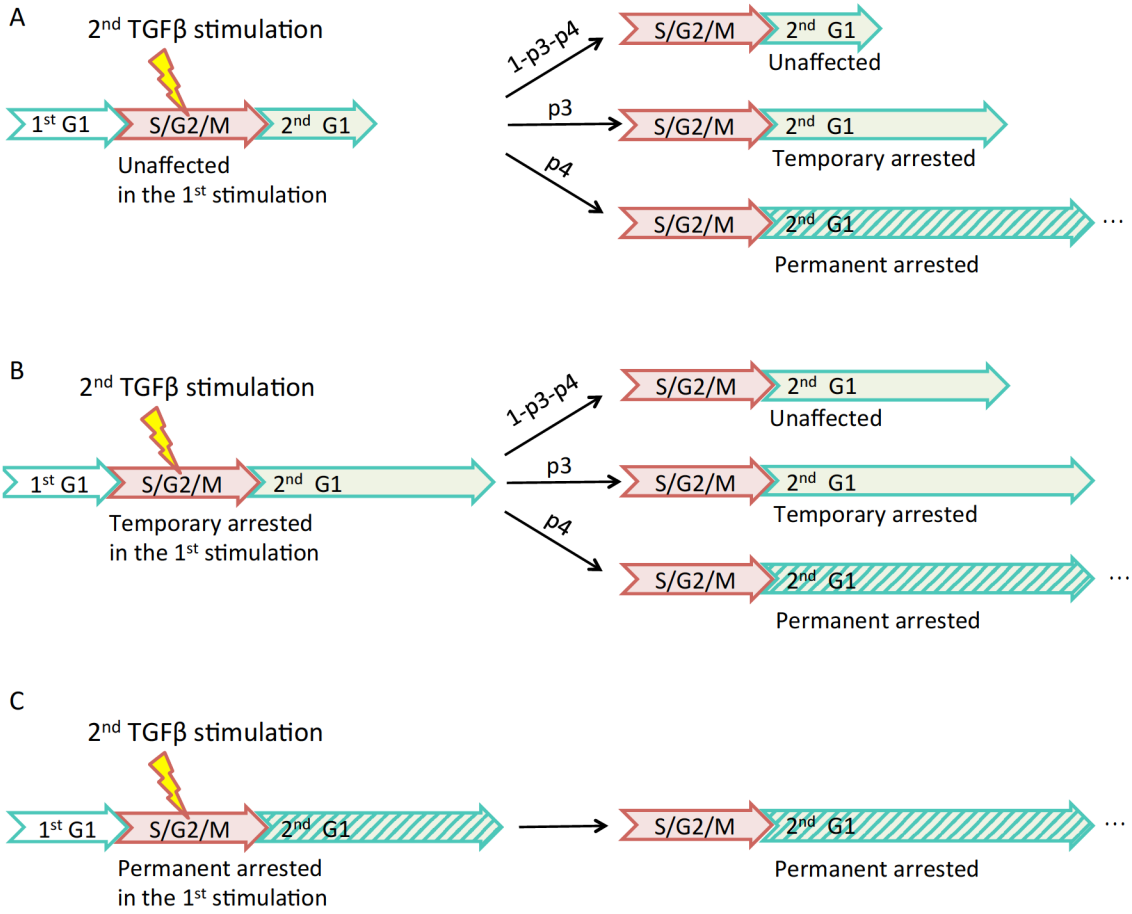


Figure 5.11 Schematic showing the assumption of cell cycle arrest in response to the second TGFβ stimulation in S/G2/M phase in the CAM. There are three groups of cells after the 1st TGFβ stimulation: unaffected, TempArrest and PermArrest in the 2nd G1 phase after cell division. (A) and (B) For unaffected cells and temporary arrested cells, they have probabilities to be unaffected, temporary arrested or permanent arrested after 2nd stimulation. After the 2nd stimulation, for the unaffected cells, they keep the same length of 2nd G1 phase as before the 2nd stimulation; for temporary arrested cells, they employ a longer 2nd G1 phase than it before stimulation; for permanent arrested cells, they stay in 2nd G1 phase. (C) For the permanent arrested cells, they stay in 2nd G1 phase.

We assumed that the cells have the same probabilities to be affected in response to every pulse of TGFβ stimulations. For the temporarily arrested cells, the G1 durations were further elongated if the cells were temporary arrested by another round of TGFβ stimulation.

Under this assumption, we simulated the cells upon double stimulations of 0, 10, 50 or 100pM TGFβ (In the pattern: 3h on, 3h off, 3h on, off). Compared with the result of single stimulation, the maximums of the fraction of G1 cells were larger upon double stimulations (Figure 5.12).

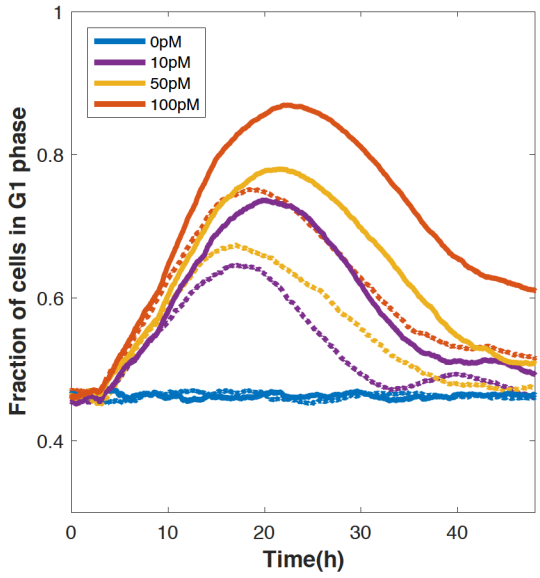


Figure 5.12 Modeling of the G1 arrest induced by single/double TGFβ stimulations. *In silico*, the cells were stimulated by the indicated dose of TGFβ. The cells were either treated by single pulse (at 0h) or double pulses (at 0h and 6h) of TGFβ stimulation. Dash lines: Single pulse of stimulation; Solid lines: Double pulses of stimulation.

Next, we would like to test the model predictions by experiments. Cell were treated in the same pattern as in the modeling: 3h on, 3h off, 3h on, off. However, the model simulations did not fit the experimental data, which implied that the assumption of responses to double stimulations was not correct (Figure 5.13).

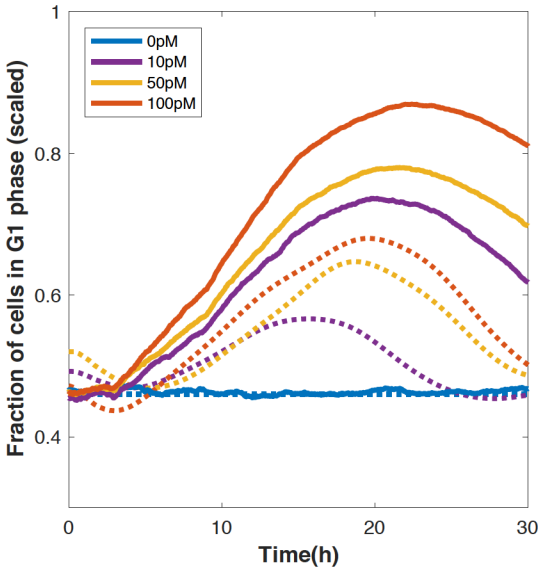


Figure 5.13 Comparison between experimental data and modeling prediction. In experiments and *in silico*, the cells were stimulated twice by the indicated doses of TGFβ (3h stimulations at 0h and 6h). Solid lines: model simulations. Dash lines: experimental data. The experimental data were scaled accordingly (Details see Chapter 9: 9.3).

5.3 Modeling suggests that cells are in a refractory state after TGFβ stimulation

In this section, we came up with another hypothesis that after one pulse of TGFβ stimulation, the cells are in a refractory state and not affected by the growth inhibitory function of another TGFβ stimulation for a certain period of time (Insensitive duration).

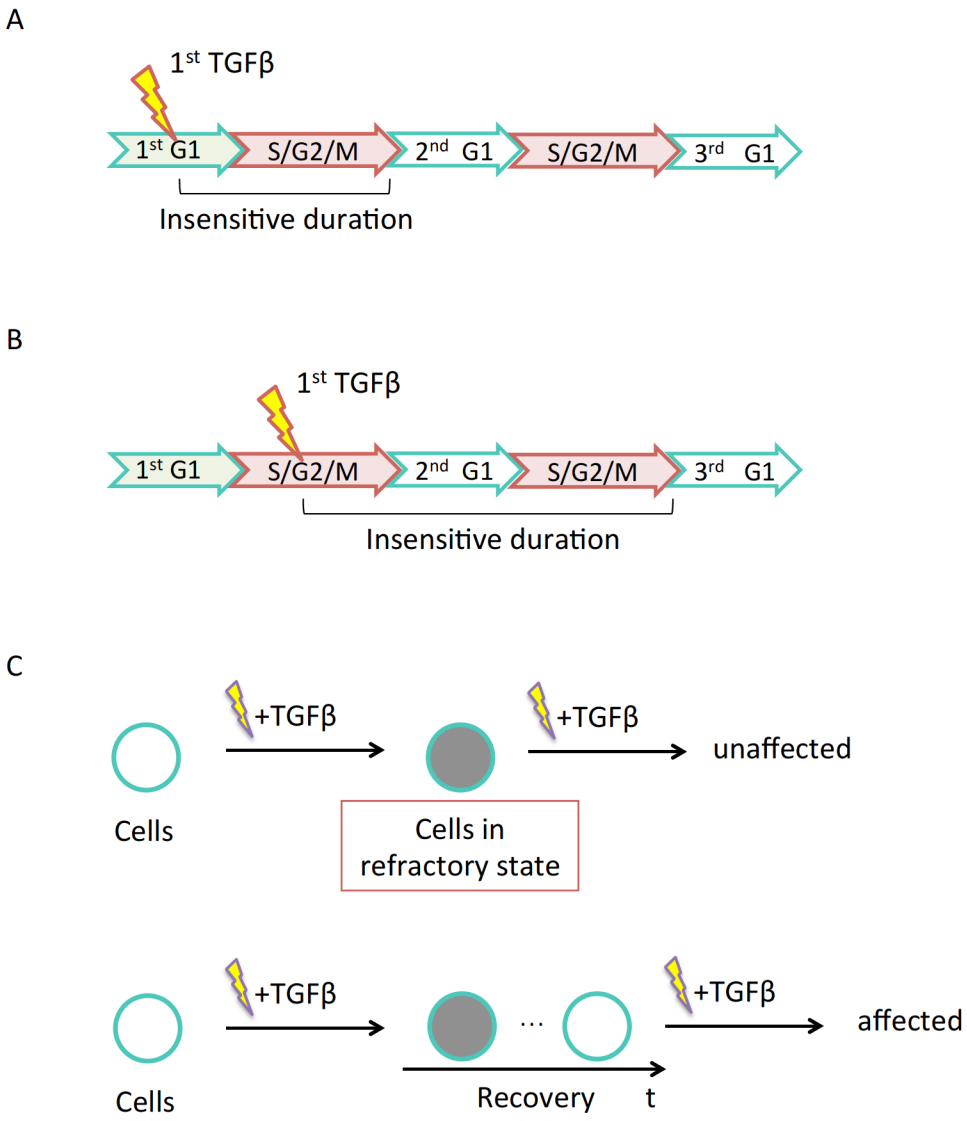


Figure 5.14 Hypothesis of cells are in a refractory state after one pulse of TGFβ stimulation. (A)Schematic showing the insensitive duration of the cells stimulated in G1 phase. (B) Schematic showing the insensitive duration of the cells stimulated in S/G2/M phase. (C) Upper panel: the cells are in a refractory state after one pulse of TGFβ stimulation, and the cells in the refractory state do not respond to TGFβ stimulation; bottom panel: the cells are in a refractory state after one pulse of TGFβ stimulation and after a period of time, the cells are not in refractory state, and respond to TGFβ stimulation.

Here we have a few assumptions. For the cells treated in G1 phase, the insensitive duration lasts from the time of stimulation to the end of M phase in this cell cycle (Figure 5.14A). For the cells treated in S/G2/M phase, as they will be affected in the next G1 phase in the following cell cycle, the insensitive duration lasts from the time of stimulation to the end of M phase in the next cell cycle (Figure 5.14B). When the cells are in the refractory state, their cell cycle progressions are not affected by TGFβ stimulations during this period. Once the cells are not in the refractory state, they respond and are arrested by TGFβ stimulation (Figure 5.14C).

Next, we would like to test our hypotheses by modeling cell cycle behaviors upon double stimulations of TGFβ under each hypothesis:

Hypothesis 1: the cells have the same probabilities to be affected in response to every pulse of TGFβ stimulations (Chapter 5:5.2 , Figure 5.10 and Figure 5.11).

Hypothesis 2: the cells are in the refractory state after the TGFβ stimulation (Figure 5.14).

Considering the cellular responses to TGFβ are dose-dependent, we simulated two different patterns of stimulations: double stimulations and step-up stimulations (Figure 5.15, Table 5-4).

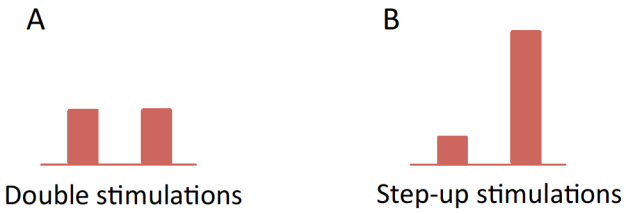


Figure 5.15 Schematic showing the patterns of double stimulations and step-up stimulations. Double stimulations: the doses of TGFβ are the same for the 1st and 2nd treatments. Step-up stimulations: the TGFβ dose in the 2nd treatment is higher than it in the 1st treatment.

Table 5-4 Models with different assumptions

	Hypothesis 1 (No refractory state)	Hypothesis 2 (Refractory state)
Double stimulations	Model 1	Model 2
Step-up stimulations	Model 3	Model 4

5.3.1 Modeling cell cycle arrest induced by double stimulations of TGFβ

In this section, we investigated the responses of cells upon double stimulations of 50pM TGFβ.

Model simulation and experimental data show that, after single pulse treatment of 50pM TGFβ, the fraction of cells in G1 phase reaches the maximum around 23h after one pulse of TGFβ stimulation (Figure 5.16A). If hypothesis 1 is true, this is the best time for the 2nd TGFβ stimulation to arrest a larger number of cells in the population. If hypothesis 2 is true, more than half of the cells (around 64.3%) are in the refractory state at this time and they will not be affected by the 2nd TGFβ stimulation. There should be observable differences between simulation results base on the two hypotheses. Thus, the time point of 2nd stimulation was set to 23h after 1st stimulation for both model simulations (Model 1 and Model 2) and experiments (Figure 5.16B). Data upon a single stimulation (with 50pM TGFβ) was used for comparisons.

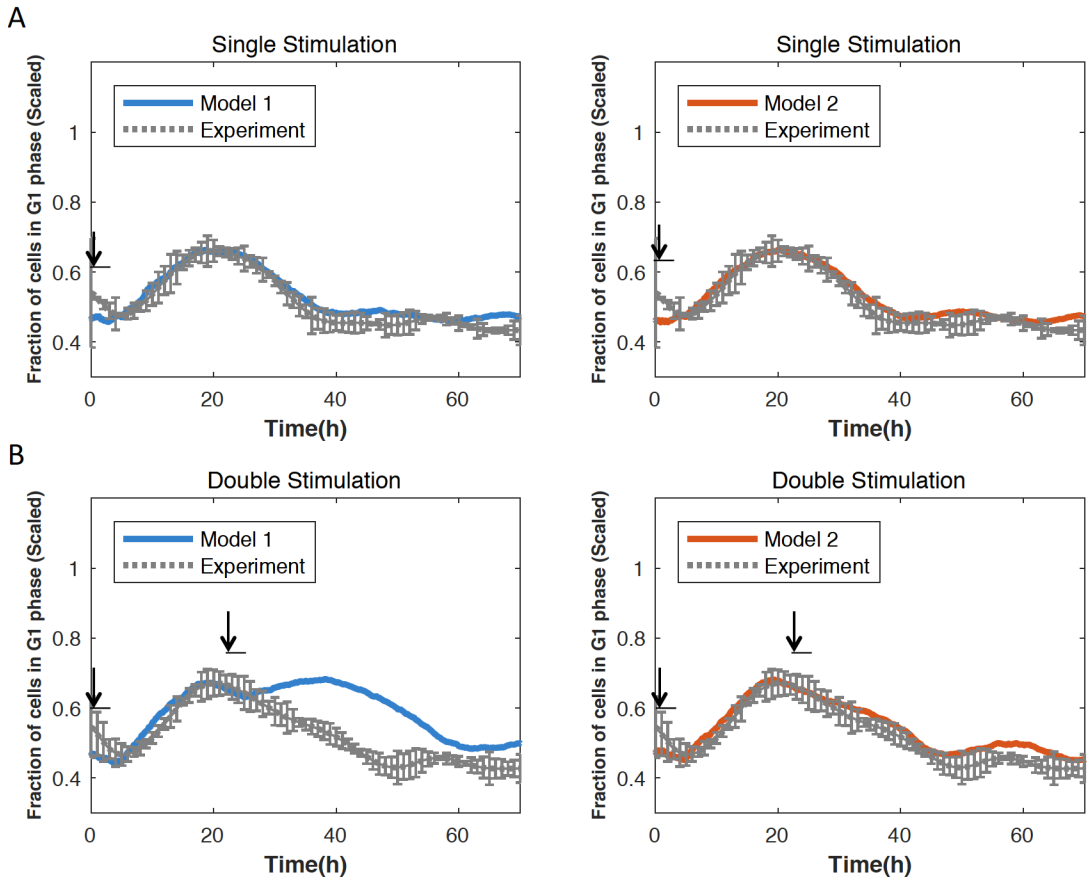


Figure 5.16 Simulations of cell-cycle arrest induced by single/double TGFβ stimulations. (A) and (B) Left: modeling based on the hypothesis that the cells had the same probabilities to be affected in response to every pulse of TGFβ stimulations. Right: modeling based on the hypothesis that the cells are in a refractory state after the TGFβ stimulation. (A) Cells were stimulated by one

(legend continued on next page)

pulse of 50pM TGFβ at 0h. The pulse lasts 3h. (B) Cells were double-stimulated by pulses of 50pM TGFβ at 0h and 23h. Each pulse lasts 3h. Black arrow with a grey line: the timing of stimulations. Solid line: model simulation. Dash line: experimental data. The experimental data were scaled accordingly (Details see Chapter 9: 9.3). The data was obtained by calculating the average values of experimental data from three replicates. Error bars: standard deviation.

Compared with the model predictions based on hypothesis 1, the experimental data showed that the fraction of cells in G1 phase was not maintained by second stimulation, suggesting that hypothesis 1 is incorrect. Model prediction based on hypothesis 2 was similar to the experimental data, suggesting the existence of a refractory behavior.

5.3.2 Modeling cell cycle arrest induced by step-up stimulations of TGFβ

We demonstrated that cells treated by 50pM TGFβ did not respond to the same dose of TGFβ. We wondered if the cells would respond to high dose of TGFβ when they were pretreated with a lower dose of TGFβ. Thus, model simulations (Model 3 and Model 4) and experiments were conducted, in which cells were stimulated with 10pM TGFβ at 0h and with 100pM TGFβ at 23h (Figure 5.17). Each pulse lasted for 3h.

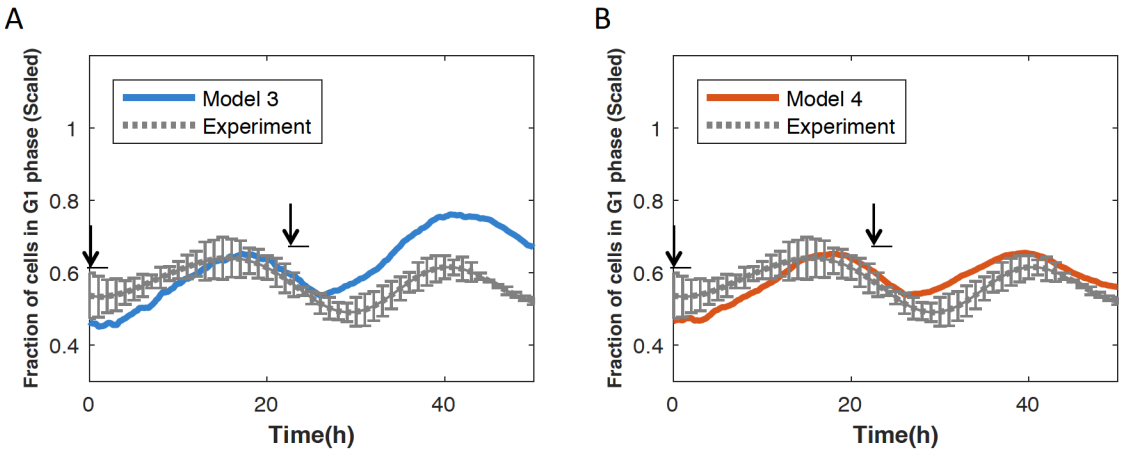


Figure 5.17 Simulations of cells stimulated with a low dose of TGFβ firstly, and then treated with a high dose of TGFβ. Cells were stimulated by one pulse of 10pM TGFβ at 0h and another pulse of 100pM TGFβ at 23h. Each pulse lasted for 3h. (A) Model3: model simulation based on the hypothesis that the cells had the same probabilities to be affected in response to every pulse of TGFβ stimulations. (B) Model4: modeling based on the hypothesis that the cells are in a refractory state after the TGFβ stimulation. Black arrow with a grey line: the timing of stimulations. Solid line: model simulation. Dash line: experimental data. The experimental data were scaled accordingly (Details see Chapter 9: 9.3). The data was obtained by calculating the average values of experimental data from three replicates. Error bars: standard deviation.

Experimental data show that the fraction of cells in G1 phase increased and reached the maximum at around 17h after 10pM TGFβ stimulation; a 2nd peak

appeared after the 2nd stimulation (100pM TGFβ) and the peak was lower than the one induced by the 1st TGFβ stimulation (10pM). However, in the model simulation based on hypothesis 1 (Model 3), the 2nd peak was much higher than the 1st one. This result suggested that the hypothesis 1 is not correct.

Model simulation data based on hypothesis 2 (Model 4) was similar to the experimental data, suggesting the hypothesis 2 is true, and low-dose treated cells would not respond to the 2nd pulse of high-dose TGFβ stimulation.

Taken together, Model 2 and Model 4 could reproduce the characteristics of experimental data, supporting the hypothesis that the cells are in the refractory state after the TGFβ stimulation.

Table 5-5 Model simulations suggest the existence of a refractory behavior

	Hypothesis 1 (No refractory state)	Hypothesis 2 (Refractory state)
Double stimulations	Model 1 ⊘	Model 2 ✓
Step-up stimulations	Model 3 ⊘	Model 4 ✓

* ⊘: the model can not represent the characteristics of experimental data; ✓: the model can represent the characteristics of experimental data.

Chapter 6

Population to Single cell framework

Although there are some software available for automatic cell tracking and analyzing, the accuracy for cell tracking in long-term live cell imaging over several generations is relatively low, especially for the cells that tend to grow in clumps. Manual correction is helpful but time-consuming. Compared with single cell data, information from the population perspective is much easier to achieve in most experiments.

Therefore, we developed a Population to Single cell framework (P2S framework) to facilitate inferring single cell information based on the easily achieved population data. We applied the P2S framework to our experimental data and obtained the cell-cycle trajectories of individual cells efficiently.

6.1 Population to Single cell framework (P2S framework)

By comparing the results of TGFβ-induced G1 arrest at the population level with those at the single cell level, we found that some interesting phenomena were hidden behind the average. For example, for cells upon 100pM TGFβ treatment, which is a saturated dose and supposed to strongly inhibit cell cycle progression, the fraction of G1 cells could reach over 80% at the population level. However, after single cell lineage tracking, we found that there were still around 40% cells escaping from the growth inhibition of TGFβ.

In asynchronously growing populations, cells are in different stages of cell cycle phases. This leads to varied responses to the same ligand treatment. The heterogeneity of cell cycle durations also influences the population behavior. Due to those complexities, it is difficult to infer single cell responses from population data. It is necessary to perform single cell analysis.

However, to obtain single cell data, on one hand, there is a high requirement of techniques and facilities in experiments for single cell analysis. For example, time-lapse fluorescence microscopy is required for live-cell imaging and high quality in every single image is essential. On the other hand, accurate single cell analysis is not easily achieved with automatic tracking software and it always needs manual correction, which is time and labor consuming.

Compared with single cell data, population-level data is much easier to acquire in most experiments. For example, it is very convenient to automatically count cells in G1 or S/G2/M phase in each image at each time point, and the results are more reliable and robust. Besides imaging, we can also get similar information by other methods (e.g. FACS).

Therefore, we decided to take the advantage of mathematical modeling and develop a Population to Single cell framework (P2S framework) to facilitate inferring single cell information based on the population data.

6.1.1 Framework workflow

To explore the cell-cycle outcomes in response to ligand treatment, we found that the key element linking the single cell and population scales is the AP parameters (Table 5-1). In this framework, we modified the PBM to describe the cell-cycle outcomes in response to TGFβ. By fitting the population data of G1 fraction changing upon stimulation, the four AP parameters (affected percentages) could be estimated. With these obtained AP parameters, single cell cycle

trajectories could then be simulated by the CAM. In this way, we could bridge the gap between population and single cell scales (Figure 6.1).

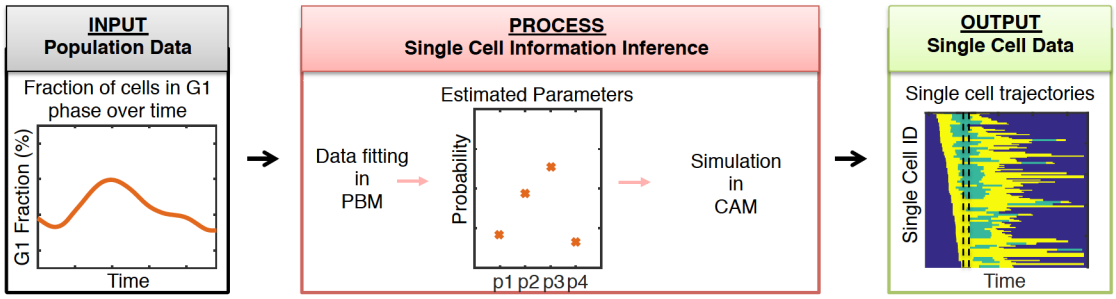


Figure 6.1 Workflow of the Population to Single cell framework (P2S framework). The framework is consist of a population balance model (PBM) and cellular automaton model (CAM), which are linked by the four ligand dose dependent parameters: p1, p2, p3 and p4. The population data (fraction of cells in G1 phase over time after ligand treatment) should be inputted into the framework. After processing, single cell trajectories will be outputted and the single cell information can be obtained by analyzing the output.

6.1.1.1 Modifying PBM for all-or-none cell-cycle outcomes in response to TGFβ stimulation

The concept and the steady state for initial condition of PBM were explained in details (Chapter 3:3.1). The PBM in Chapter 3 described the unaffected group and TempArrest group induced by TGFβ. Here, we further developed the model to incorporate PermArrest group and fit the all-or-none G1 arrest response induced by TGFβ.

For the cells permanently arrested in G1 phase, they were not progress. For the cells in S/G2/M phase and marked as permanently arrested, they only progressed when they were in S/G2/M stage and then were permanently arrested in G1 phase after cell division. We used two sets of ODEs to describe cell progression of TempArrest group (Equation 24 and 25).

$$\frac{dG_i^{pA}}{dt} = 0 \tag{24}$$

$$\frac{dM_j^{pA}}{dt} = \begin{cases} -\frac{M_j^{pA}}{\tau}\left(1 - p_{M\to G}(j)\right) - M_j^{pA} \cdot p_{M\to G}(j) & for\ j = 1 \\ \frac{M_{j-1}^{pA}}{\tau}\left(1 - p_{M\to G}(j-1)\right) - \frac{M_j^{pA}}{\tau}\left(1 - p_{M\to G}(j)\right) - M_j^{pA} \cdot p_{M\to G}(j) & for\ j = 2, \dots, m \end{cases} \tag{25}$$

where G_i^{pA} and M_j^{pA} are the cell numbers (PermArrest) in the i-th sub-phase in G1 and j-th sub-phase in S/G2/M, respectively; τ is the duration of the sub-phase, which is the same as in unaffected group; $p_{M\to G}(j)$ is the phase transition

probabilities of the j–th sub-phase in S/G2/M. Because cells do not arrest at S/G2/M phase, the transition probability $p_{M\to G}$ is the same as in unaffected group.

The equations in PBM incorporating TGFβ-induced cell cycle arrest are summarized (Table 6-1).

Table 6-1 Summary of the equations in PBM in P2S framework

Untreated and unaffected cells

$$\frac{dG_i}{dt}=\begin{cases} 2\frac{M_m}{\tau}\left(1-p_{M\to G}(m)\right)+2\sum_{j=1}^{m}M_j\cdot p_{M\to G}(j)-\frac{G_i}{\tau}\left(1-p_{G\to M}(i)\right)-G_i\cdot p_{G\to M}(i) & \\ & for\ i=1 \\ \frac{G_{i-1}}{\tau}\left(1-p_{G\to M}(i-1)\right)-\frac{G_i}{\tau}\left(1-p_{G\to M}(i)\right)-G_i\cdot p_{G\to M}(i) & \\ & for\ i=2,\dots,n \end{cases}$$

$$\frac{dM_j}{dt}=\begin{cases} \frac{G_n}{\tau}\left(1-p_{G\to M}(n)\right)+\sum_{i=1}^{n}G_i\cdot p_{G\to M}(i)-\frac{M_j}{\tau}\left(1-p_{M\to G}(j)\right)-M_j\cdot p_{M\to G}(j) & \\ +\frac{G'_{n'}}{\tau}\left(1-p'_{G\to M}(n')\right)+\sum_{i=1}^{n'}G'_i\cdot p'_{G\to M}(i) & for\ j=1 \\ \frac{M_{j-1}}{\tau}\left(1-p_{M\to G}(j-1)\right)-\frac{M_j}{\tau}\left(1-p_{M\to G}(j)\right)-M_j\cdot p_{M\to G}(j) & for\ j=2,\dots,m \end{cases}$$

Temporary arrested cells

$$\frac{dG'_i}{dt}=\begin{cases} 2\frac{M'_m}{\tau}\left(1-p_{M\to G}(m)\right)+2\sum_{j=1}^{m}M'_j\cdot p_{M\to G}(j)-\frac{G'_i}{\tau}\left(1-p'_{G\to M}(i)\right)-G'_i\cdot p'_{G\to M}(i) & \\ & for\ i=1 \\ \frac{G'_{i-1}}{\tau}\left(1-p'_{G\to M}(i-1)\right)-\frac{G'_i}{\tau}\left(1-p'_{G\to M}(i)\right)-G'_i\cdot p'_{G\to M}(i) & for\ i=2,\dots,n \end{cases}$$

$$\frac{dM'_j}{dt}=\begin{cases} -\frac{M'_j}{\tau}\left(1-p_{M\to G}(j)\right)-M'_j\cdot p_{M\to G}(j) & for\ j=1 \\ \frac{M'_{j-1}}{\tau}\left(1-p_{M\to G}(j-1)\right)-\frac{M'_j}{\tau}\left(1-p_{M\to G}(j)\right)-M'_j\cdot p_{M\to G}(j) & for\ j=2,\dots,m \end{cases}$$

Permanently arrested cells

$$\frac{dG_i^{pA}}{dt} = 0$$

$$\frac{dM_j^{pA}}{dt} = \begin{cases} -\frac{M_j^{pA}}{\tau}\left(1 - p_{M\to G}(j)\right) - M_j^{pA} \cdot p_{M\to G}(j) & for\ j = 1 \\ \frac{M_{j-1}^{pA}}{\tau}\left(1 - p_{M\to G}(j-1)\right) - \frac{M_j^{pA}}{\tau}\left(1 - p_{M\to G}(j)\right) - M_j^{pA} \cdot p_{M\to G}(j) & for\ j = 2, \dots, m \end{cases}$$

To incorporate the effect of TGFβ-induced all-or-none G1 arrest into the PBM model, as in CAM (Table 5-2), we employed the four AP parameters: p1, p2, p3 and p4. Upon TGFβ stimulation, cells were categorized into eight groups by applying the four AP parameters to the number of cells in each sub-phases at steady state (Table 6-2).

Table 6-2 Initial conditions of PBM in P2S framework

Variable	Stimulated@phase	Arrest state	Initial Conditions
G_i	-	unaffected	$G_i^{steady} \cdot (1 - p1 - p2)$
G_i'	G1	TempArrest	$G_i^{steady} \cdot p1$
	S/G2/M	TempArrest	0
G_i^{pA}	G1	PermArrest	$G_i^{steady} \cdot p2$
	S/G2/M	PermArrest	0
M_j	-	unaffected	$M_j^{steady} \cdot (1 - p3 - p4)$
M_j'	S/G2/M	TempArrest	$M_j^{steady} \cdot p3$
M_j^{pA}	S/G2/M	PermArrest	$M_j^{steady} \cdot p4$

The values of parameters used in PBM were set based on experimental data (Table 6-3). The values of parameters 1-12 are the same as in Table 3-3 and Table 5-2. The values of parameter $p'_{G\to M}(i)$ were calculated based on the values of n′, μ'_G and σ'_G in the similar way as calculating $p_{G\to M}(i)$ in Chapter 3: 3.1.3 . The values of parameter $p'_{G\to M}(i)$ were not shown for space limitation.

Table 6-3 Summary of the derived parameter values based on experimental data

	Parameter	Annotation	Value and Unit
1	τ	The duration of the subphase	2.6213 min
2	n	The number of subphases in G1 phase	550
3	m	The number of subphases in S/G2/M phase	637
4	n′	The number of subphases in arrested G1 phase	1746
5	μ_G	Parameter μ and σ of the log-normally distributed G1 phase duration	6.1646 min
6	σ_G		0.2218 min
7	μ_M	Parameter μ and σ of the log-normally distributed S/G2/M phase duration	6.6681 min
8	σ_M		0.1505 min
9	μ'_G	Parameter μ and σ of the log-normally distributed TempArrest G1 phase duration	7.3673 min
10	σ'_G		0.2123 min
11	$p_{G\to M}(i)$	The probability of cells transiting into S/G2/M phase at i-th sub-phase in G1 phase; i=1, … , n	(See Appendix Chapter 9: 9.1)
12	$p_{M\to G}(j)$	The probability of cells transiting into G1 phase at j-th sub-phase in S/G2/M phase; j=1, … , m	(See Appendix Chapter 9: 9.1)
13	$p'_{G\to M}(i)$	The probability of cells transiting into S/G2/M phase at i-th sub-phase in arrested G1 phase; i=1, … , n’	-

6.1.1.2 Single Cell Information (AP parameters) Inference

The AP parameters are important for reflecting and evaluating the cytostatic effect of TGFβ treatment. However, they were obtained by accurate single cell tracking, which is time-consuming. In contrast, the G1 fraction (the fraction of cells in G1 phase) dynamic data is at the population level, and could be easily obtained by automatic image analysis. In this framework, we proposed a way to infer the AP parameters through fitting the experimental data (G1 fraction profile) in PBM.

Although the sets of ODEs could not be analytically solved, it is still possible to perform parameter estimation for the AP parameters by some software or toolboxes, e.g. Data2Dynamics[141] and PottersWheel[142]. However, due to the huge number of ODEs in the PBM, it would take a long time to obtain the results. Thus we estimated the values of AP parameters by parameter searching. For each time of data fitting, we run more than 10000 simulations by using Latin hypercube sampling method [143,144], which is an efficient method to generate random samples

for parameter values and spread samples evenly across all possible values[145]. For the four AP parameters, 50000 parameter sets were uniformly sampled in the 4-dimensional space during the range of [0, 1] for all the 4 parameters. After that, two constrains were applied: 0<=p1+p2<=1; 0<=p3+p4<=1 and the sets that did not match the constrains were removed (around 12500 sets were left). G1 fractions dynamics were simulated based on the values of each set of AP parameters. The Euclidean distance ($d\ (S, E)$) between stimulated results (S_t) and experimental data (E_t) were calculated (Equation 26, Figure 6.2).

$$d(S,E) = \sqrt{(S_1 - E_1)^2 + (S_2 - E_2)^2 + \cdots + (S_k - E_k)^2} \qquad k \in \{1, \ldots, t\} \qquad (26)$$

where $d(S,E)$ is the Euclidean distance between stimulated results and experimental data; S_k is the simulated G1 fraction at time point k; E_k is the experimental G1 fraction at time point k.

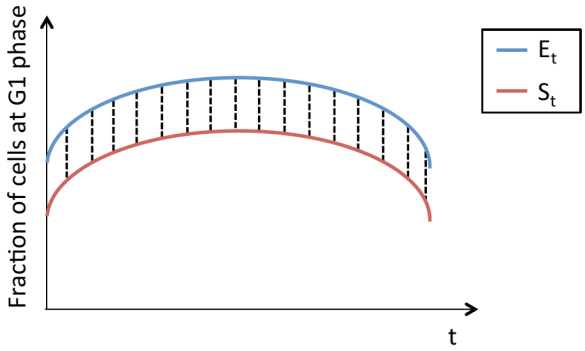


Figure 6.2 Schematic showing that the difference between stimulation and experimental data. Euclidean distance $d\ (S, E)$ between stimulated results (S_t) and experimental data (E_t), were calculated based on equation 26.

The parameter sets were then sorted by the corresponding Euclidean distance $d\ (S, E)$. The parameter set with the minimum Euclidean distance was chosen. By performing data fitting independently for 10 times and averaging the results, we got the values of fitted parameters. With these fitted parameters from PBM, single cell cycle trajectories could then be simulated by CAM (the detail of the CAM is in Chapter 5: 5.1).

To confirm the robustness of this method, we artificially assign a set of values for the AP parameters and simulated a G1 fraction profile (Figure 6.3A, black dash line), then added 10% noise to it (Figure 6.3C, red line). By performing data fitting, we got the values of fitted parameters, which were closed to the values that we initially set before (Figure 6.3 B and D), suggesting the robustness of the method.

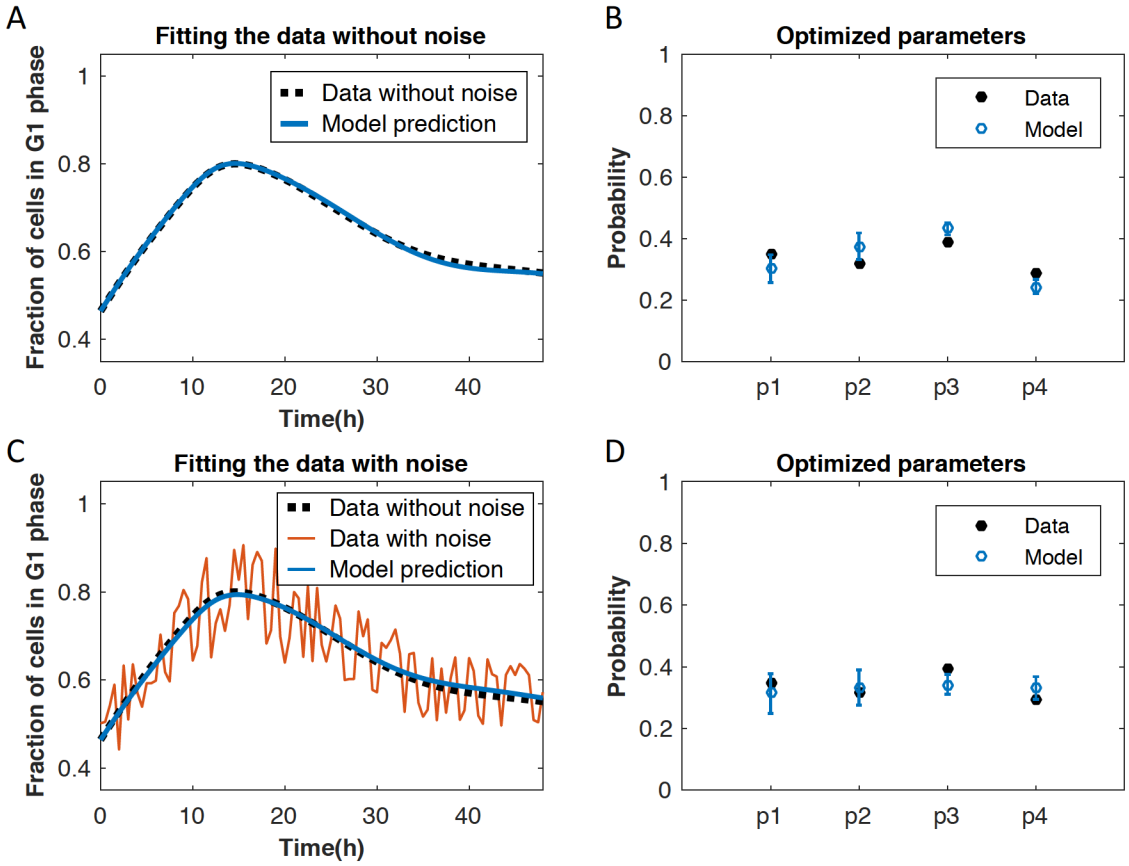


Figure 6.3 The method of parameter estimation is robust. (A) And (B) Pseudodata of G1 fraction dynamics was generated by artificially setting the four AP parameters (p1=0.35, p2=0.32, p3=0.39, p4=0.29) and running simulation (A, black dash line). The fitted parameters (B, blue crosses) were shown in comparison with the originally set ones (B, black circles). Error bar: stand deviation. Simulation results based on the fitted parameters (A, blue lines) were plotted against the pseudodata. (C) and (D) Pseudodata was generated as in (A), with ±10% noise introduced (C, red line). The fitted parameters and the original values were shown in (D).

6.1.2 Framework validation

To test the applicability of this framework for experimental data, we used the data in Figure 2.8A as input, and estimated the four AP parameters for each TGFβ dose independently (Figure 6.4).

Notably, there is a delay of around 3h before the increase of G1 fraction after TGFβ treatment. We only use the experimental data after the initiation of G1 fraction increase for data fitting.

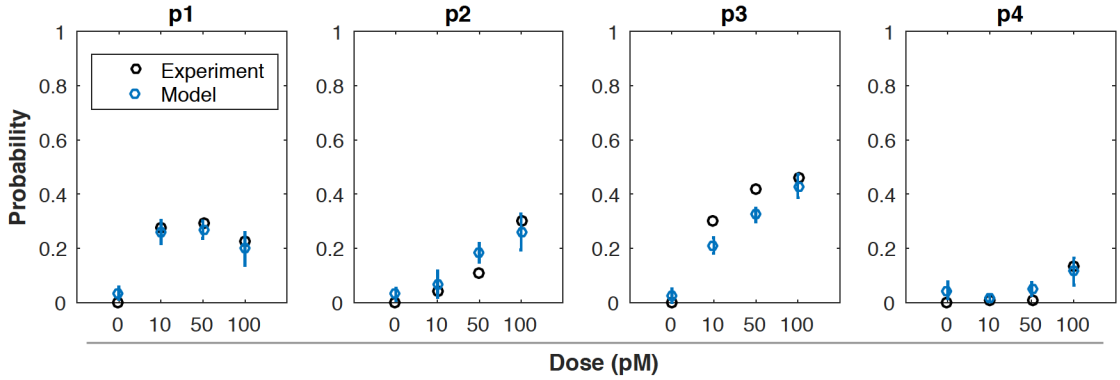


Figure 6.4 Comparison between the experimental data and model estimation. Using the data in Figure 2.8A as input in the P2S framework, values for the four AP parameters were estimated for each TGFβ doses (blue crosses); error bar: stand deviation. Values calculated from the experimental data (shown in Figure 4.4B) were also shown here for comparison (black circles). p1 (probability of TempArrest when stimulated at G1 phase), p2 (probability of PermArrest when stimulated at G1 phase), p3 (probability of TempArrest in the following G1 phase when stimulated at S/G2/M phase) and p4 (probability of PermArrest in the following G1 phase when stimulated at S/G2/M phase).

Most optimized values of the AP parameter were similar to the experimental measurements. One value of p2 (50pM) and two values of p3 (10pM, 50pM) were not very close to the experimental data. There were two main reasons for the deviation: firstly, we assumed that the population has reached the steady state before TGFβ treatment in the model. However, in experiments, the population might slightly deviate from the steady state. Secondly, the random generated samples for searching might not include identical values to the experimental data. Thus we only obtained the approximate values.

Although the estimated values did not exactly match the experimental data, the dose-dependent patterns were similar. With increasing concentrations of TGFβ stimulation, the value of p1 slightly decreased and the values of p2, p3 and p4 increased.

Numerous single cell cycle trajectories were further simulated (Figure 6.5), and the results were plotted in the same way of the corresponding experimental data (Figure 4.1 and Figure 4.4). Due to method of Latin hypercube sampling, we only got approximate values. Thus when using the data of control (0pM treatment) as input, the estimated values of the four AP parameters were approximate to zero, but not zero. Therefore, there could be some arrested cells in the simulation for the control group (Figure 6.5 and Figure 6.6).

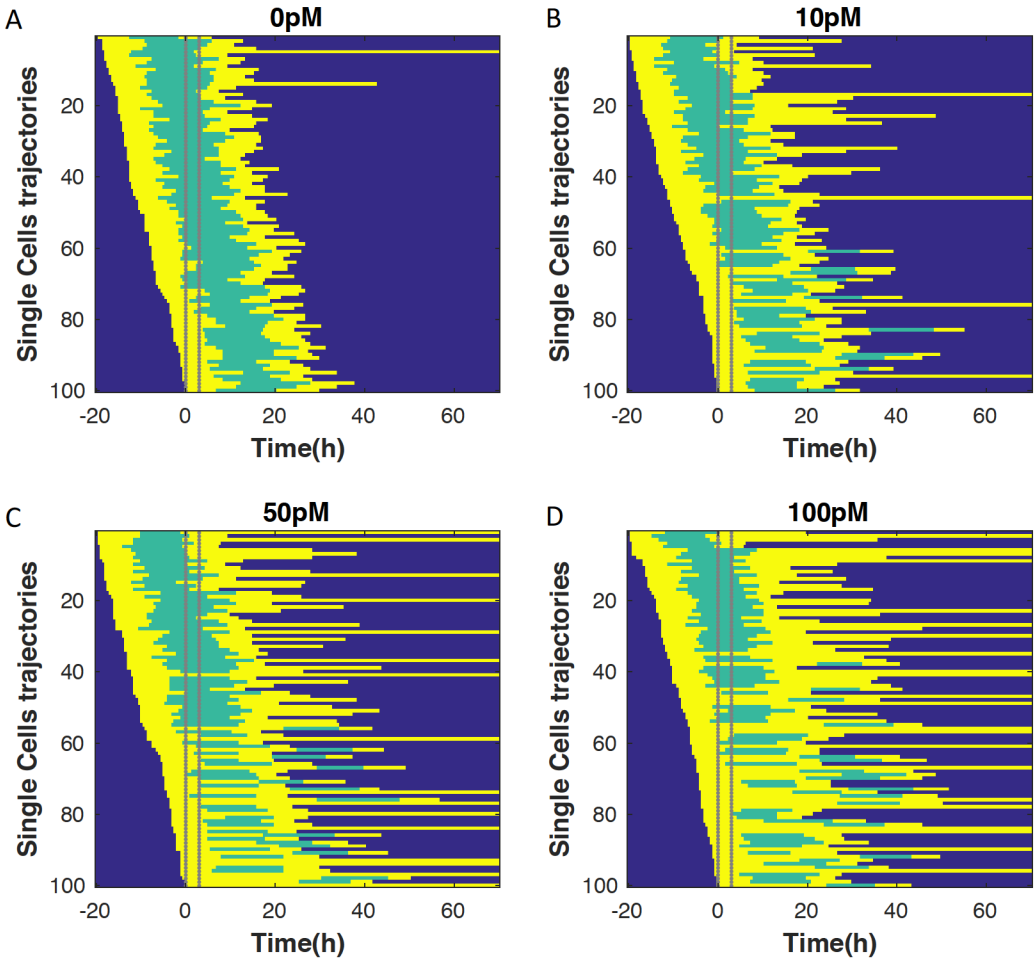


Figure 6.5 Simulations of single cell trajectories in P2S framework. (A)-(D) Using the data in Figure 2.8A as input in the P2S framework, the output of single cell trajectories were simulated to mimic the asynchronously cycling cells upon 0, 10, 50 or 100pM TGFβ1 treatment. The time of adding TGFβ was artificially set as zero. The duration of TGFβ treatment was shown (Within dash lines). Each row represents the trajectory of an individual cell. Yellow color indicates the cell was in G1 phase and the green color indicates the cell was in S/G2/M phase. Cells were sorted by the time of birth. 100 cells were randomly chosen from the group upon each dose of TGFβ treatment.

The durations of 1st or 2nd G1 phase of the stimulated cells were calculated and compared with the experimental data (Figure 6.6A and B). The fractions of cells in G1 phase were also calculated for each time point (Figure 6.6C).

The simulated single cell data corresponded well with the experimental quantification. In addition, the dynamics of G1 fraction were also consistent between simulation and experimental measurement (Figure 6.6C). These results verified the applicability of the framework.

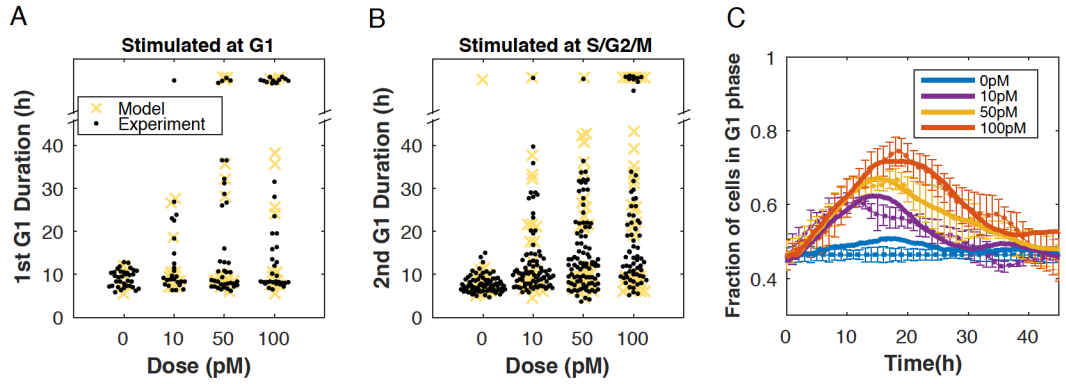


Figure 6.6 Simulations from P2S framework correspond to experimental data. (A) and (B) The 1st G1 and 2nd G1 phase duration of the cells in Figure 6.5 were analyzed quantitatively. 250 cells were randomly chosen. (A) The 1st G1 duration of the cells stimulated at G1 phase; (B) The 2nd G1 duration of the cells stimulated at S/G2/M phase; black dots: model simulation, yellow cross: experimental data; 1st G1: the G1 phase in the current cell cycle upon TGFβ treatment; 2nd G1: the G1 phase in the next cell cycle following the cell cycle upon TGFβ treatment. (C) According to the data in Figure 6.5, the fractions of cells in G1 phase over time were calculated. Experimental data used as the input (dash lines) and model simulations obtained as output (solid lines) were shown.

6.2 Application of P2S framework

We performed a series of experiments, in which the growing cells were treated with 0, 5, 10, 20, 50, 100, 200 and 400pM TGFβ and imaged under the microscope for 60h. Due to the low efficiency of manual analysis, we only quantitatively analyzed hundreds of cells upon 0, 10, 50 and 100pM TGFβ stimulation (control, low dose, medium dose and high dose).

We would like to use the P2S framework to find out the fractions of arrested cells in each group (cells treated by 0, 5, 10, 20, 50, 100, 200 and 400pM TGFβ).

6.2.1 Estimating the fractions of arrested cells in response to varied doses of TGFβ

The G1 fraction over time of each group were automatically quantified by the software CellProfiler[146,147]. The time-course data of G1 fraction in each group were used as input for the P2S framework. Then the four AP parameters for the affected percentages were estimated (Figure 6.7). With these obtained AP parameters, single cell cycle trajectories could then be simulated by the CAM.

The results showed that the fraction of temporary arrested cells did not always increase with elevating doses of TGFβ treatments. Instead, the fraction of TempArrest dropped down while the fraction of PermArrest increased. More and

more cells shifted into PermArrest group in response to increasing doses of TGFβ treatments.

Based on the curves of each AP parameter, we could also predict the values of AP parameters for other doses. For example, we did not measure the data of cells upon 75pM TGFβ treatment experimentally but we could estimate the approximate values of p1, p2, p3 and p4 from the curve, which are 0.23, 0.22, 0.38 and 0.08, respectively.

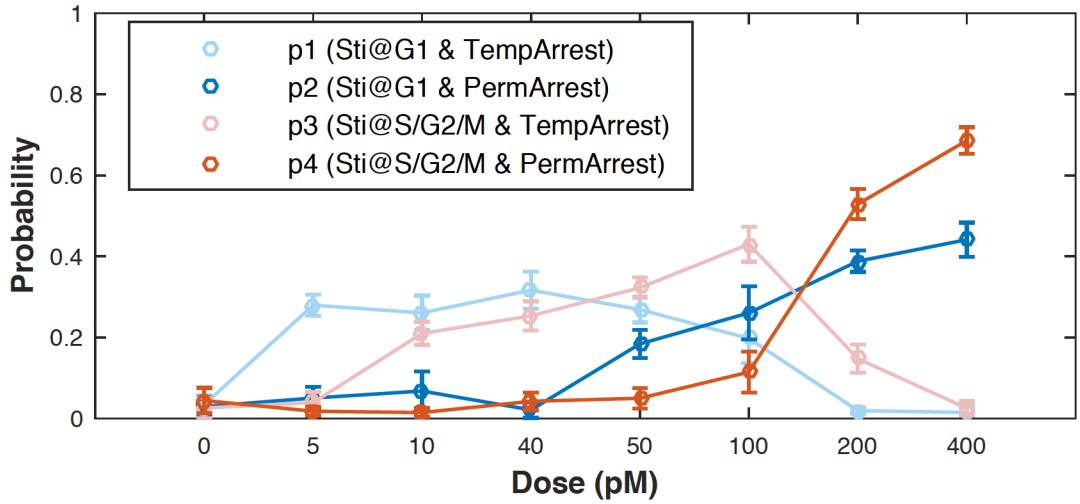


Figure 6.7 Affected fractions of cells upon ligand treatment were estimated by P2S framework. The fractions of affected cells (equal to the value of p1, p2, p3 and p4) upon increasing doses of TGFβ stimulation (0, 5, 10, 20, 50, 100, 200 and 400pM) were estimated by the P2S framework. p1 (probability of TempArrest when stimulated at G1 phase), p2 (probability of PermArrest when stimulated at G1 phase), p3 (probability of TempArrest in the following G1 phase when stimulated at S/G2/M phase) and p4 (probability of PermArrest in the following G1 phase when stimulated at S/G2/M phase). Light blue: p1; dark blue: p2; light red: p3, dark red: p4. Error bar: stand deviation.

6.2.2 Model predictions inspired the exploration of TGFβ-induced mitotic failure

Interestingly, when exposed to 400pM ligand, p1 and p3 (probabilities of TempArrest) dropped down to 0. The cell seems to make binary decisions upon 400pM TGFβ stimulation, either totally unaffected or permanent arrested in G1 phase.

We looked back into the imaging data and found that nearly 40% cells had undergone mitotic failure when stimulated at G1 phase. These cells exited the mitotic cell cycle and differentiated into polyploid cells with giant nuclei, termed "endoreduplication"[148,149]. Though they did not divide after the S/G2/M phase, the

signal of FUCCI reporter vanished, therefore they were categorized as being permanent arrested by the model (Figure 6.8).

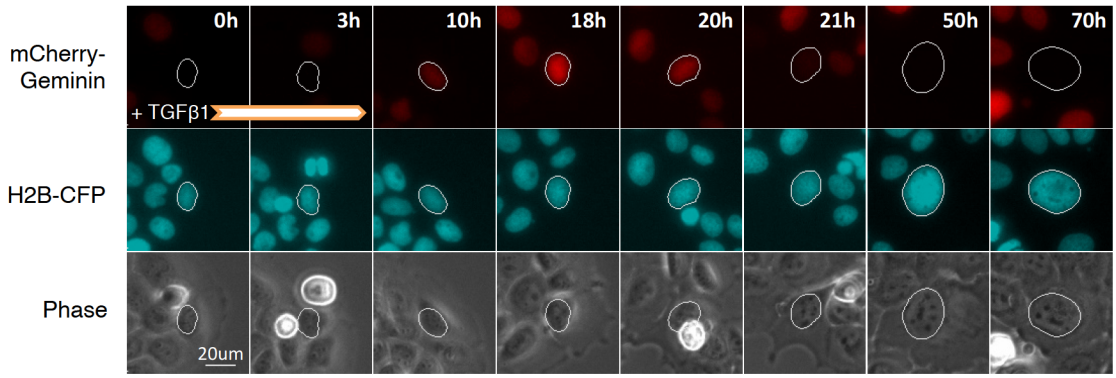


Figure 6.8 Cell undergoes mitotic failure. An example of a cell underwent mitotic failure when stimulated at G1 phase. The nuclear segmentation of the cell was indicated as white circle. There was no nuclear mCherry-Geminin signal during 0-3h, which suggested this cell was in G1 phase when stimulated by TGFβ. There was an observable nuclear mCherry-Geminin signal from 10h-20h, which indicated that the cell progressed into S/G2/M phase. The cell did not divide but the nuclear mCherry-Geminin signal vanished. The cells were treated by 200pM TGFβ1 during 0-3h.

The cells underwent mitotic failure were quantitatively analyzed (Figure 6.9). For the cell stimulated at G1 phase, mitotic failure was observed in the cells upon TGFβ treatment over 100pM. With increasing concentration of TGFβ, more cells shifted into the group of mitotic failure.

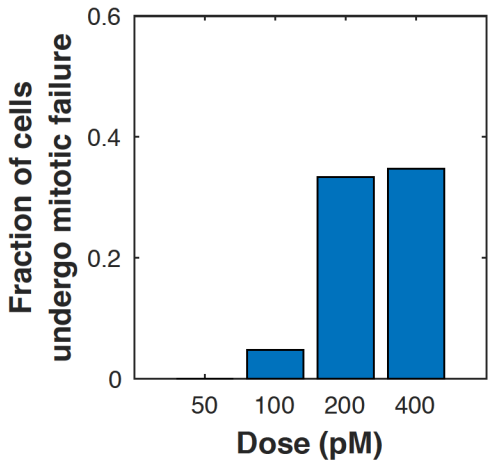


Figure 6.9 Mitotic failure induced by high dose of TGFβ treatment. The fractions of cells underwent mitotic failure upon 50, 100, 200 and 400pM TGFβ1 stimulation at G1 phase.

Notably, the mitotic failure only happened for treatment at G1 phase, we did not observe it in cells that were stimulated at S/G2/M phase. This result also supports that cells in G1 phase were more sensitive to TGFβ in terms of cell cycle alternation.

Chapter 7

Summary and discussion

Single cell analysis can reveal hidden information behind the population average. In this thesis, using time-lapse microscopy, we investigated cell cycle arrest induced by TGFβ in individual, asynchronously growing cells.

Combining mathematical modeling with experimental measurements, our results indicate that TGFβ stimulation triggers immediate or heritable all-or-none G1 arrest response. The fraction of affected cells increases along with the elevating ligand concentrations. The sensitivity to TGFβ is both cell cycle phase and ligand dose dependent. Considering the importance of single cell information and the difficulties in the single cell analysis, we developed a P2S framework to infer single cell information efficiently from population dynamics.

In this chapter, we summarize the project, discuss possible molecular mechanisms underlying the observations and give an outlook for future applications.

7.1 Summary of the work

In biology, population-averaged assays play an essential role in studies of complex signaling and transcriptional networks. In recent decades, investigations on heterogeneous cellular behaviors unveil meaningful biological information masked by population-averaged measurements[150-153].

Due to the lack of techniques, previous studies on the growth inhibitory function of TGFβ were conducted at the population level, overlooking the cellular heterogeneity (Figure 7.1A). Therefore, although population behaviors were well investigated, single cell responses to TGFβ were left unexplored.

In this project, a fluorescence system containing reporters for both TGFβ signaling and cell cycle was constructed. By time-lapse fluorescence live-cell microscopy, we made it possible to visualize cell cycle phases and TGFβ/Smad signaling in individual cells. Various dimensions of single-cell measurements were taken into consideration in the investigation.

Firstly, responses of cells in different cell cycle phases (G1 or S/G2/M phase) upon TGFβ stimulation were quantified (Figure 7.1B, Chapter 3: 3.2.2.1 and Chapter 4: 4.1). The experimental design of single pulse stimulation facilitated the studies on phase-dependent response. Besides immediate G1 arrest induced by TGFβ stimulation at G1 phase, we found the inheritable, delayed G1 arrest induced by stimulation during S/G2/M phase. This phenomenon suggests that, the whole cell cycle should be subjected to TGFβ in order to achieve maximal cell growth inhibitory effect.

Secondly, the heterogeneous G1 durations after TGFβ treatment were quantitatively analyzed (Figure 7.1C, Chapter 3: 3.2.2.2 and Chapter 4: 4.2). Our results demonstrate that TGFβ stimulation triggers all-or-none G1 arrest response. Importantly, this observation implies that there are always some cells escaping the growth control of TGFβ.

Thirdly, the cell-cycle outcomes were inspected based on the timing between and within cell cycle phases of TGFβ stimulation (Chapter 4: 4.3 and 4.4).

The sensitivity of the cells to the concentration of TGFβ is phase-dependent (Figure 7.1D, Chapter 4: 4.3). Cells in G1 phase are most sensitive to TGFβ in terms of cell cycle alternation: TGFβ induces a larger fraction of cells undergoing permanent arrest or mitotic failure in G1 than it in S/G2/M phase. Cells in S phase are least sensitive: they do not respond to TGFβ below 20pM, while cells in other phases respond. These phase-specific responses to TGFβ are probably determined

by the phase-dependent protein abundance, since protein levels determine the cell-specific response to TGFβ[90].

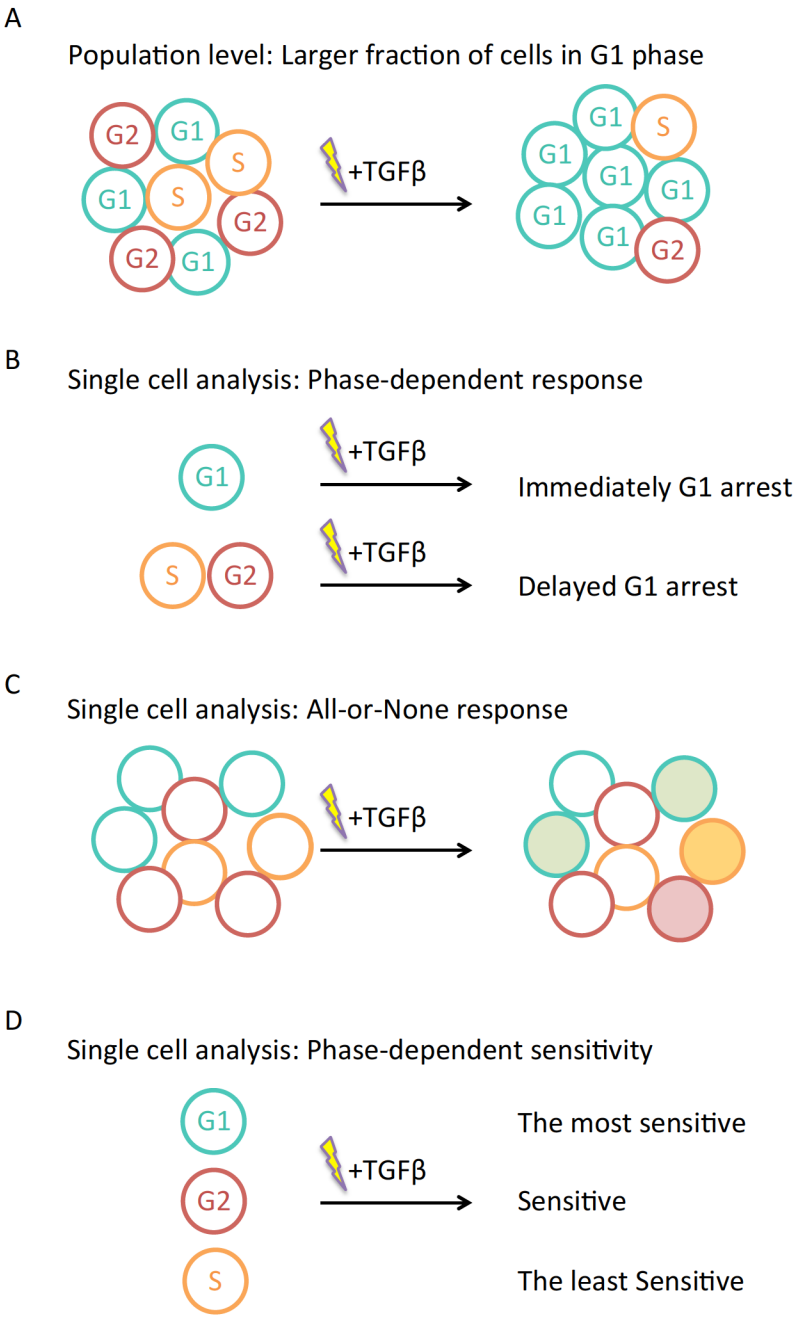


Figure 7.1 Single cells analysis unveils the heterogeneous cell-cycle arrest response induced by TGFβ. (A) Studies of TGFβ-induced G1 arrest at population level. (B), (C) and (D) Studies of TGFβ-induced G1 arrest at single cell level in this project. (B) TGFβ induces immediate G1 arrest for cells at G1 phase and inheritable, delayed G1 arrest in the next cell cycle for cells in S/G2/M phase. (C) TGFβ induces all-or-none G1 arrest response: some cells are unaffected and some cells are arrested. (D) The sensitivity of the cells to the concentration of TGFβ is Phase-dependent: the cells in G1 phase are the most sensitive and the cells in S phase are the least sensitive. Circles: cells in G1, S or G2/M phase are indicated as G1, S, G2, respectively.

The percentage of affected cells is dependent on the timing within the phase upon high dose of TGFβ stimulation (Chapter 4: 4.4). In the early G1 phase, 50pM TGFβ induced less arrested cells than it in middle/late G1 phase.

However, the lengthened G1 durations were not dependent on the stages of cell cycle phases. This finding agrees with the knowledge that the G1 restriction point or the key decision window is located at the end of G1 or G2/M phase, respectively. Cells integrate signals during the individual cell cycle phase, and make decisions just before entering the next phase. Moreover, it also suggests robust regulation by TGFβ on cell cycle within each phase, which guarantees precisely control in tissue hemostasis.

7.2 Significance of all-or-none response

Cells sense extracellular stimuli and integrate the information to respond. At the population level, the responses are typically dose dependent, which represent analogue signal information processing of stimulus[154,155]. However, there are increasing evidences that various signaling pathways respond in a digital manner at the single cell level[156-159], i.e. all-or-none response.

Studies of T cells revealed that, on one hand, digital signaling enables a rapid and sensitive response to stimulation exceeding the threshold; on the other hand, it prevents spurious T cell activation under weak stimulations[158]. In neurons, digital signaling could integrate multiple inputs into all-or-none action potentials[160]. The TGFβ-induced all-or-none G1 arrest response could be critical for defining sharp threshold of stimulations. It prevents cells being affected by weak stimulation (noise from other signaling pathways) while preserving their high sensitivity to strong stimulations. Moreover, the single cells could be “on” or “off” in response to the same dose of TGFβ treatment and the “off” cells are able to escape from cytostatic effect of TGFβ. This could allow some cells to survive from the cytostatic effect.

Notably, although the cells were treated with 3h TGFβ in this work, the G1 arrest response could last for 30h -70h. How the cells ‘remember’ a transient differentiation stimulus has been investigated for decades. To answer this question, Monnd and Jacob proposed six possible types of signal transduction circuits, which were variations of a double-negative feedback circuit[161]. Besides double-negative feedback circuits, the self-sustaining patterns (gene expression or protein activity) could also be achieved by positive feedback circuits[162]. Those systems can toggle between two stable states (‘’on” or “off”) but would not rest in

intermediate states. Thus the all-or-none response could be the result of a self-sustaining signaling activation system.

7.3 Molecular mechanism of the all-or-none G1 arrest

In epithelial cells, one of the most important TGFβ-induced cytostatic effects is implemented by the activation of *CDKN1A*, which encodes p21[9]. The p21 halts the cell in G1 phase by directly inhibiting the activity of the CyclinE-CDK2 complex. In turn, the CyclinE-CDK2 complex drives the G1/S transition and leads to the degradation of p21. A double-negative feedback loop is formed[163,164]. The immediate all-or-none response in G1 phase triggered by TGFβ could be explained by the bistability of the p21-CDK2 system (Figure 7.2).

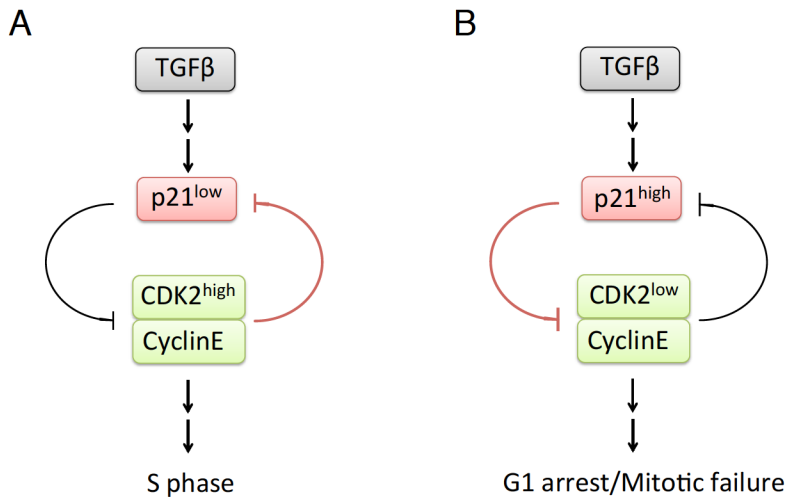


Figure 7.2 Schematic diagram of the double-negative feedback loop in p21 and CDK2-CyclinE regulatory system. (A) Low p21 and high CDK2 level triggers G1/S transition. (B) High p21 and low CDK2 level maintains the state of G1 arrest or triggers mitotic failure. Red blunted arrow: dominant inhibition.

TGFβ induces heterogeneously elevated levels of p21 in the cell population (Chapter 4: 4.6). The cells with low p21 levels escape from the growth inhibitory effect of TGFβ and progress to S phase. The cells with higher p21 levels have a higher probability to be arrested. And the onset of p21 decay and activation of CDK2 lead to the complete p21 degradation[163]. Then the cells recover from arrest and enter the S phase. In this way, the double-negative feedback loop involving p21 and CDK2 generates bistable steady states: $p21^{low}$/$CDK2^{high}$ associated cycling state, and $p21^{high}$/$CDK2^{low}$ associated quiescent state[163,165,166].

Besides the immediate G1 arrest, the heritable G1 arrest triggered by TGFβ could also be explained by the p21-CDK2 system.

There is a bifurcation point at the mitotic exit where cells make a decision between two different fates (Figure 7.3). Cells with intermediate levels of CDK2 continue to build up the CDK2 activity and progress in the next cell cycle ($CDK2^{inc}$ fate). Other cells with low CDK2 levels enter a transient quiescent state ($CDK2^{low}$ fate). This bifurcation of cell fate is controlled by p21[165].

For the cells stimulated by TGFβ at S/G2/M phase, varied levels of p21 are induced. When a cell arrives at the restriction window (at the end of cell cycle), it integrates the mitogenic signals and chooses one path ($p21^{low}/CDK2^{inc}$ fate or $p21^{high}/CDK2^{low}$ fate) upon completion of mitosis. Then the cell will be unaffected ($CDK2^{inc}$) or arrested ($CDK2^{low}$) in the next G1 phase after cell division.

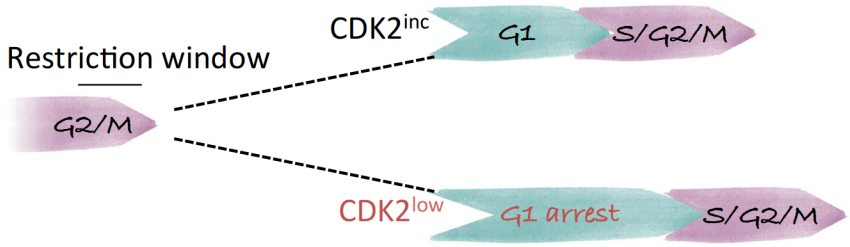


Figure 7.3 Schematic showing the key decision point of the cell-cycle commitment at the mitotic exit. The cells choose one fate between $CDK2^{inc}$ and $CDK2^{low}$ in the restriction window at the end of previous cell cycle. $CDK2^{inc}$: the cells with intermediate levels of CDK2 activity continue to build up the CDK2 activity and progress in the next cell cycle; $CDK2^{low}$: the cells with low levels of CDK2 activity enter a transient quiescent state.

7.4 Refractory behavior of TGFβ signaling pathway

We found that the cells were in a refractory state after one pulse of TGFβ stimulation and would not be affected by the further TGFβ stimulations. The molecular mechanism underlying this refractory behavior is not clear.

Previous studies on TGFβ signaling pathway focused on TGFβ receptors and Smads signaling. The degradation of TGFβ receptors, which is mediated by endocytosis, plays a critical role in turning off the signaling[167]. The half-life of cell surface receptors was reported to be between 45min to 4h, depending on the specific cell lines[168-171]. Continuous trafficking of receptors between endosomes and plasma membrane is required for TGFβ signaling[172,173]. When all the available competent receptors were occupied, the cell would be unable to perceive extracellular ligand[174]. It was observed that cells were in a refractory state after 1h TGFβ treatment and did not respond to further acute stimulation (the Smad2 phosphorylation was not induced)[174]. 12-24h after ligand removal, the cells

completely recovered their capacity to respond to TGFβ (Smad2 phosphorylation could be induced)[174].

However, although the receptors recovered at 12h[174], we found that the cytostatic response to TGFβ was not recovered even at 23h after ligand removal (Chapter 5:5.3). Thus there exist regulations on other signaling reactions than the ligand-receptor binding that result in the refractory behavior of cytostatic response. The molecular mechanism remains to be unraveled. Some non-canonical signaling pathways (Smad independent) or negative feedbacks could be involved.

7.5 Comparison of TGFβ and DNA damage induced cell-cycle arrest

It should be noted that p21 is also a downstream target in p53 signaling, which is activated by DNA damage. Although transcriptional activation of p21 and cell-cycle arrest are induced by both DNA damage and TGFβ, TGFβ specifically induces p21 through a p53-independent mechanism[175], which distinguishes the cell cycle control by TGFβ from that by stress. There are several characteristics distinguish the cell cycle arrest induced by TGFβ from that induced by DNA damage (Table 7-1).

Table 7-1 Comparison of cell cycle arrest induced by TGFβ or DNA damage

	TGFβ	DNA damage
Aim	To keep homeostasis of the population[9]	To repair the DNA in individual cells[176]
Cell-cycle outcomes	G1 arrest	G1, S and G2/M arrest[177]
Most sensitive phase	G1 phase	G2 phase[177]
Cell in refractory state after one stimulation	Yes	-

Firstly, the TGFβ-induced cell-cycle arrest plays an essential role in keeping homeostasis of the population[9]. While the DNA damage delays the cell-cycle progression to repair DNA in individual cells[176]. Secondly, cells treated with TGFβ are arrested in G1 phase. While cells are arrested immediately in either G1, S or G2/M phase by DNA damage[177]. Thirdly, cells in G1 phase are most sensitive to TGFβ. While cells in G2 phase show the highest sensitivity to DNA damage[177]. Fourthly, after one pulse of TGFβ stimulation, the cells are in a refractory state and are not affected by another pulse of TGFβ treatment during a period over one or two cell cycle phases (Figure 5.14). While DNA damage always activates checkpoints in each cell cycle phase and delays cell-cycle progression.

In the future studies, investigation on the upstream TGFβ signaling is crucial to unravel the molecular mechanism of TGFβ-induced cell-cycle arrest.

7.6 Model prediction guides experimental design

We employed PBM and CAM to describe the cell cycle and validate some hypotheses on TGFβ-induced G1 arrest. Combining model prediction and experimental confirmation, we found several crucial characteristics of the TGFβ-induced G1 arrest.

Model predictions of inheritable G1 arrest inspired us to investigate the phase durations over two cell cycles (Chapter 4: 4.1). In addition, model prediction of insensitive duration inspired us to investigate the strategy of cells to escape from the growth inhibitory effect of TGFβ (Chapter 5:5.3).

Modeling not only inspires us to design experiments to test model predictions, but also provides new perspectives to grasp the information from the data.

From experiments, we know that the addition of TGFβ induces a larger fraction of cells in G1 phase and four groups of arrested cells (Stimulated at G1 or S/G2/M phase: TempArrest or PermArrest). But it is difficult to determine the contribution of each group to the temporal dynamics of G1 fraction. However, we can easily explore it by mathematical modeling (Figure 7.4).

The result of parameter scanning indicates that the increase of G1 fractions in 10h after stimulation is mainly caused by arrested cells that are stimulated in G1 phase (p1 and p2, Figure 7.4A and B), while the peak at 20h is mainly the result of arrested cells that are treated in S/G2/M phase (p3 and p4, Figure 7.4C and D).

In addition, mathematical modeling strongly supports the existence of inheritable, delayed G1 arrest. Assuming that cells at S/G2/M phase do not respond, we set p3 and p4 to zeros. When varying the p1 or p2, G1 fraction in the simulation always drops back to the initial level in 20h, even for conditions that all G1 cells are permanently arrested (i.e. p2=1 in Figure 7.4B). However, experimental results show that the G1 fraction remains at high levels at 20h for TGFβ concentrations over 5pM (Figure 2.8A). This higher level of G1 fraction is caused by the portion of cells stimulated at S/G2/M phase, which is confirmed by simulations in which p3 or p4 was varied, leading to maximal G1 fraction at 20h (Figure 7.4C and D). These results verified the existence of inheritable, delayed G1 arrest from the modeling perspective.

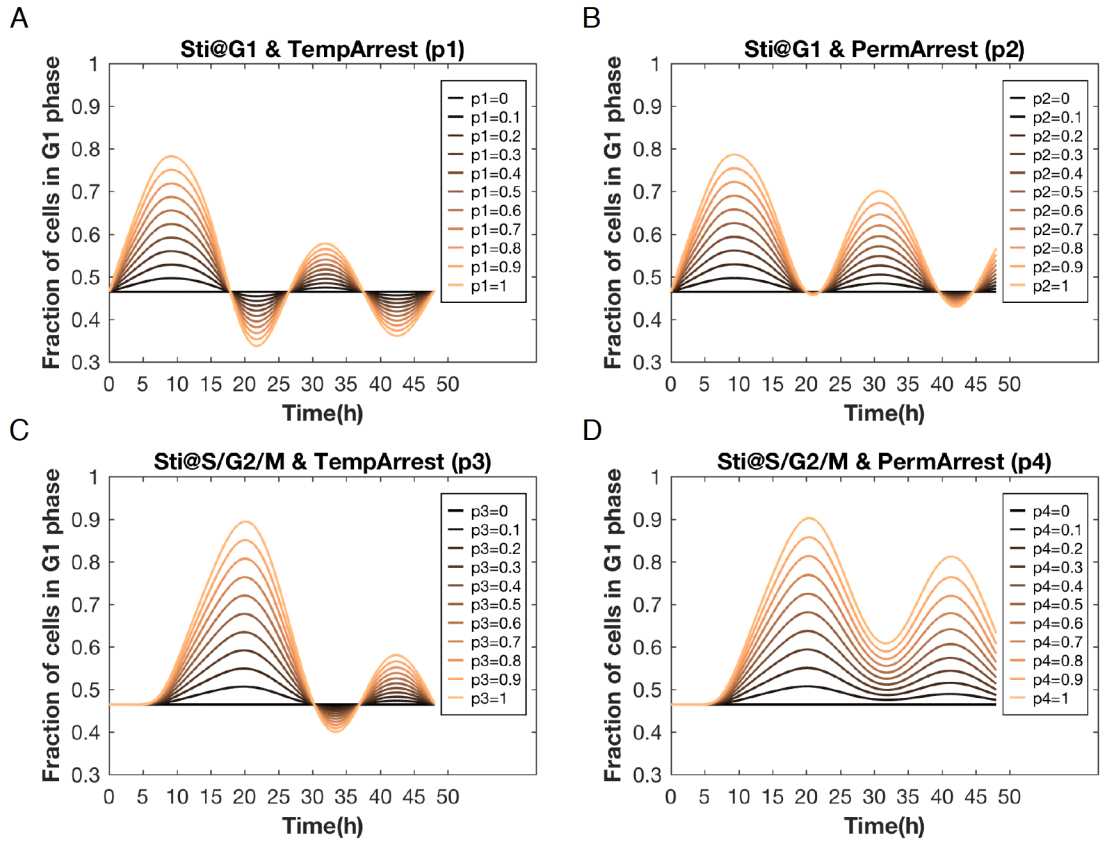


Figure 7.4 Parameter scanning results for the four AP parameters (TGFβ dose dependent) in PBM. The default values of p1-p4 are 0. Each of p1 (A), p2 (B), p3 (C) and p4 (D) was scanned between 0 and 1 independently; simulation results were recorded and the G1 fraction over time were calculated respectively.

What's more, the simulations also provide reference on choosing proper time-points for experimental measurement. Knowing the doubling time of the cells, we could predict the ligand-induced oscillation of G1 fraction. When design experiments, we could choose the proper time points to measure samples according to the model predictions. For instance, if we would like to screen mutants that interrupt/enhance the effect of TGFβ-induced G1 arrest, we could take the samples around 10h or 20h after ligand treatment to measure fractions of cells in G1 phase, which show the maximum level of G1 fraction caused by arrested cells stimulated in G1 (Figure 7.4A and B) or S/G2/M phase (Figure 7.4C and D), respectively.

7.7 Outlook

In this study, we provide a P2S framework to facilitate exploring hidden information in the TGFβ-induced cytostatic response. The framework consist of a

population balance model (PBM) and a cellular automaton model (CAM), which are linked by four AP parameters: p1, p2, p3 and p4.

The P2S framework can be used to facilitate future studies on TGFβ signaling. For instance, we can explore novel factors in TGFβ signaling pathway (Figure 7.5).

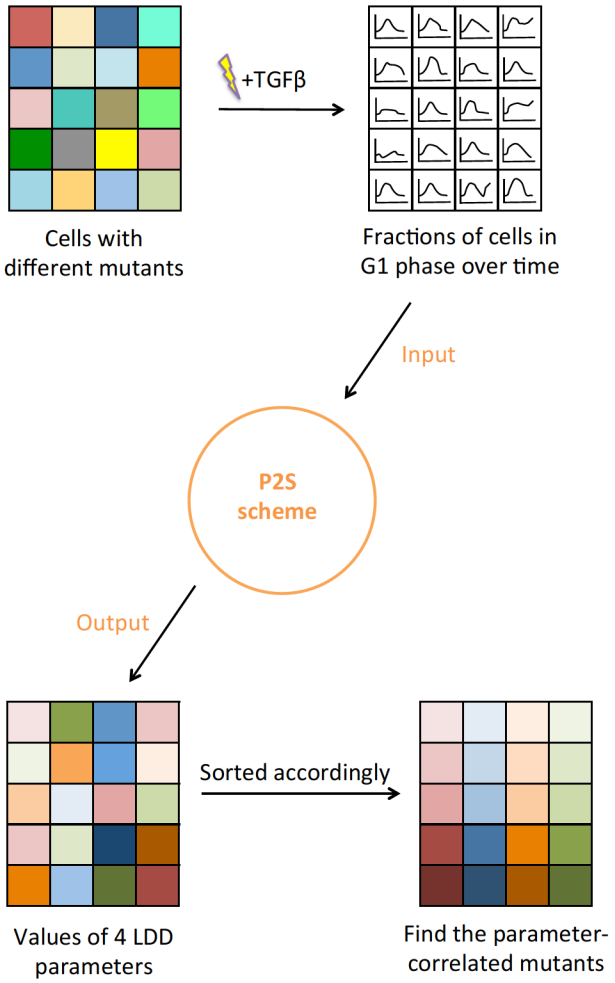


Figure 7.5 Schematic showing the workflow of exploring novel factors in TGFβ signaling pathway. Cells with different mutants/knock-down proteins/overexpressed proteins -> Treat the cells with TGFβ -> Monitor the cells under the microscope -> Automatically calculate the fractions of cells in G1 phase over time -> Estimate the values of 4 AP parameters -> Find the parameter-correlated genes/proteins.

Firstly, a library of cells with different mutants/knock-down proteins/overexpressed proteins should be built. Then we can generate their temporal dynamics of G1 fraction upon TGFβ stimulation and analyze it by the P2S framework to obtain the values of the AP parameters. Then the mutants/proteins related to significant changes (decreasing/increasing) of AP parameter values can be identified.

By using the P2S framework, we could not only identify the genes related to TGFβ-induced G1 arrest, but also distinguish them by their specific targets, e.g., we could find out phase-dependent genes/proteins networks targeted by TGFβ, based on the p1/p2-related genes and p3/p4-related genes identified by the P2S framework, which are involved in G1 and S/G2/M phase dependent response, respectively.

Further, this framework not only benefits studies on TGFβ signaling, but can also be generalized to other cytostatic factors, and potential for high-throughput cytostatic drug screening. Moreover, by modifying the P2S and incorporating protein-protein interacting networks, it is possible to describe and even predict the dynamics of molecular events (e.g. activities of CDKs, circadian genes, stress response factors) during the cell cycle progression, providing a powerful analysis tool for many basic/applied researches.

Altogether, this work brings novel insights into the TGFβ-induced cell cycle arrest and provides tools to investigate single cell behavior, benefiting future studies on cytostatic responses, as well as development of therapeutical strategies for cell proliferation control.

Chapter 8

Materials and methods

8.1 Cell culture

HaCaT cells were cultured in Dulbecco's Modified Eagle Medium (Gibco, Catalog #21969-035) with 10% fetal bovine serum (Merck Biochrom, Catalog #S0615), 2 mM L-Glutamine (Gibco, Catalog #25030-024), 100 units/mL penicillin and 100 µg/mL streptomycin (Gibco, Catalog #15140-122) at 37 °C and in 5% CO2.

HEK293T cells were cultured in Dulbecco's Modified Eagle Medium (Lonza, Catalog #BE12– 707F) with 10% fetal bovine serum (Merck Biochrom, Catalog #S0615), 2 mM L-Glutamine (Gibco, Catalog #25030–024), 100 units/mL penicillin and 100 µg/mL streptomycin (Gibco, Catalog #15140–122) at 37 °C and 5% CO2.

8.2 Plasmid preparation

The plasmids of pCSII-EF-mCherry-hGeminin(1/110) is a gift from Prof. Atsushi Miyawaki(Brain Science Institute, RIKEN). pREX-H2B-ECFP-IRES-BSD was constructed by Yuchao Li. To generate pREX-H2B-ECFP-IRES-BSD, the EYFP-Smad2 on the pREX-EYFP-Smad2-IRES-BSD (a gift from Prof. Xuedong Liu, University of Colorado Boulder) was replaced with the H2B-ECFP on the pCMV-H2B-ECFP (a gift from Dr. Jianying Yang, University of Freiburg). Plasmids for lentivirus production (pCMV-VSV-G, pMDLg/pRRE and pRSV-Rev) and retrovirus production (pCL-Ampho) were all kindly provided by Prof. Dr. Xuedong Liu.

8.3 Retrovirus and Lentivirus production

To generate retrovirus for stably transducing the H2B-ECFP, HEK293T cells for Packaging were plated onto 60 mm dishes (Greiner Bio-One, Catalog #628160) one day before being cotransfected with pREX-H2B-ECFP-IRES-BSD and pCL-Ampho using Lipofectamine 3000(Invitrogen) following the manufacturer's protocol. 48 hours after transfection, medium containing retroviral particles was filtered through 0.45µm syringe filters (Carl Roth, Catalog #XH43.1). Lentivirus for mCherry-hGeminin (1/110) transduction were prepared in the same way by cotransfecting the pCSII-EF-mCherry-hGeminin(1/110) plasmid with lentiviral packaging and envelope plasmids (pCMV-VSV-G, pMDLg/pRRE and pRSV-Rev). Viral supernatant was used for transduction immediately, or kept in −80 °C freezer for long-term storage.

8.4 Stable cell line construction

The stable EYFP-Smad2 expressing HaCaT cell line was obtained from the lab of Prof. Xuedong Liu (University of Colorado Boulder). We generated stable HaCaT cell lines co-expressing EYFP-Smad2, H2B-ECFP and mCherry-hGeminin by sequential transduction of the above retrovirus and lentivirus. For viral transduction, target cells were plated onto 100 mm dishes (Greiner Bio-One, Catalog #664160). After overnight incubation, transduction was performed by applying the viral supernatant supplemented with 8 μg/mL of Polybrene (Merck Chemicals, Catalog #TR-1003-G). Single colony isolation was performed by flow cytometry sorting and confirmed by live cell imaging.

8.5 Fluorescence-activated cell sorting (FACS)

Cultured cells were trypsinized and suspended at a concentration of ~ 1 million cells per 1 ml in Dulbecco's Modified Eagle Medium lacking phenol red (Gibco, Catalog #31053-028), containing 10% fetal bovine serum (Merck Biochrom, Catalog #S0615), 2 mM L-Glutamine (Gibco, Catalog #25030-024), 100 units/mL penicillin and 100 μg/mL streptomycin (Gibco, Catalog #15140-122). All the samples were measured by BD FACS Aria II flow cytometer (BD Biocsiences, San Jose, USA). 458nm laser and 473/10 filter were used for ECFP, 488nm laser and 542/27 filter were used for EYFP, 561nm laser and 615/25 filter were used for mCherry. The data was analyzed by the FACSDivaTM software (Version 6.1.3).

When sorting cells, Plasmocin ™ Prophylactic (InvivoGen, Catalog #ant-mpp) was supplied at a final concentration of 5 μg/ml for the prevention of mycoplasma contamination.

8.6 DRAQ5 staining

Cells were seeded in the T75 flask (greiner bio-one, Catalog #658175) and incubated overnight. Samples were taken out of the incubator after indicated stimulation conditions and washed by DPBS (Gibco, Catalog #14190-094) for three times followed by treatment of Tris-EDTA buffer (Sigma-aldrich, Catalog #93283) for 10min. After that, cells were trysinized by trypsin-EDTA (gibco, Catalog #25300) for 3 min and suspended at a concentration of 4×10^5 cells/mL in Dulbecco's Modified Eagle Medium lacking phenol red (Gibco, Catalog #31053-028), containing 10% fetal bovine serum (Merck Biochrom, Catalog #S0615), 2 mM L-Glutamine (Gibco, Catalog #25030-024), 100 units/mL penicillin and 100 μg/mL streptomycin (Gibco, Catalog #15140-122). The samples were incubated

with 40µM DRAQ5 solution (BD biosciences, Catalog #564903) for 30min at room temperature. After that, cells were analyzed directly without further washing.

Cells were analyzed using BD FACS Aria II flow cytometer (BD Biocsiences, San Jose, USA). mCherry was excited by a 561 nm laser line, and DRAQ5 was excited by a 633 nm laser line. Fluorescence signals were collected at 615 nm (615/25) for mCherry and at 710 nm (710/ 50) for DRAQ5. The data were analyzed using FlowJo software (Version 10.4.1).

8.7 Immunofluorescence

Cells were seeded in the µ-Slide 8 Well (ibidi, Catalog #80826) and incubated overnight. Samples were taken out of the incubator after indicated stimulation conditions and washed by DPBS (Gibco, Catalog #14190-094) for three times followed by fixation in 4% PFA for 10 minutes. Permeabilization was performed with pre-chilled 100% methanol for 10 minutes after another DPBS washing (three times). Samples were blocked in PBS Fish Gel (ROCKLAND, Catalog #MB-066-0100) overnight at 4°C. After that, samples were incubated with primary antibody for 1h at room temperature, washed by DPBS three times, and then incubated with secondary antibody for 1h at room temperature. The primary antibody used in this work: p21 Waf1/Cip1 (12D1) primary antibody (Cell Signaling, Catalog #2947T) with a dilution of 1:400. The Secondary antibody used was the Anti-Rabbit IgG (H+L), F(ab')2 Fragment (Alexa Fluor® 647 Conjugate) (Cell Signaling, Catalog # 4414, dilution: 1:500). After being washed by DPBS three times, the samples were imaged in mounting medium (ibidi, Catalog # 50001).

Images were acquired using a Plan-Apochromat 40 x/0.95 Korr Ph 3 M27 objective (Zeiss, Catalog # 420661-9970) mounted on a Zeiss Axio-Observer Z1 fluorescence microscope. The data was analyzed automatically in the free open-source software CellProfiler (Version 3.0.0)[146,147].

8.8 Cell lysate preparation and western blot

HaCaT cells were lysed in RIPA buffer (Cell Signaling, Catalog #9806), supplemented with 1 mM PMSF, 1 mM NaF, protease inhibitor cocktail (Roche, Catalog #04 693 132 001) and phosphatase inhibitor cocktail (Roche, Catalog #04 906 845 001). Protein concentrations were determined using Pierce BCA protein assay kit (Thermo Fisher Scientific, Catalog #23225) according to the manufacturer's instructions.

Equal amount of proteins were mixed with 4x Protein Sample Loading Buffer (LI-COR Biosciences, Catalog # 928-40004). The proteins were separated by SDS–PAGE gels (Bio-Rad, Catalog #456-1045) and transferred onto nitrocellulose membrane. The membrane were incubated with primary antibody overnight at 4°C, then washed three times by TBST (Santa Cruz, Catalog #sc-24953) and incubated with secondary antibody for 1h at room temperature. The following primary antibodies were used in this work (Table 8-1):

Table 8-1 Summary of primary antibodies used in western blot

Primary Antibody	Provider	Catalog	Dilution
Phospho-Smad2 (Ser465/467)	Cell Signaling	#3108	1:500
Smad2	Cell Signaling	#3103	1:1000
Phospho-Smad3 (Ser423/425)	Cell Signaling	#9520	1:1000
Smad3	Cell Signaling	#9523	1:1000
Smad4	Cell Signaling	#9515	1:1000
β-Actin	Cell Signaling	#3700	1:10000
Lamin A/C	Cell Signaling	#4777	1:2000
GRB2	Cell Signaling	#3972	1:1000

Secondary antibodies used for western blot were anti-rabbit IgG (H+L) DyLight 800 4X PEG Conjugate (Cell Signaling, Catalog #5151, dilution: 1:15000) and anti-mouse IgG (H+L) DyLight 680 Conjugate (Cell Signaling, Catalog #5470, dilution: 1:15000). Western blot images were acquired using Odyssey CLx Imaging System (LI-COR Biosciences, Catalog # 9140).

8.9 Nuclear and cytoplasmic protein extraction

Cells were seeded in the 10cm dished (Greiner CELLSTAR, Catalog #664160) and incubated overnight. The cells were harvested in ice-cold DPBS (Gibco, Catalog #14190-094) supplemented with protease inhibitor cocktail (Roche, Catalog #04 693 132 001). Nuclear and cytoplasmic proteins were extracted using NE-PER™ Nuclear and Cytoplasmic Extraction Reagents (Thermo Fisher Scientific, Catalog #23225) according to the manufacturer's instructions. The samples were further analyzed by western blot.

8.10 Live cell microscopy

Prior to microscopy, cells were cultured in µ-Slide 8 Well (ibidi, Catalog #80826) with Dulbecco's Modified Eagle Medium lacking phenol red (Gibco, Catalog #31053-028), containing 10% fetal bovine serum (Merck Biochrom, Catalog #S0615), 2 mM L-Glutamine (Gibco, Catalog #25030-024), 100 units/mL penicillin and 100 µg/mL streptomycin (Gibco, Catalog #15140-122).

Cell imaging experiments were performed using a Zeiss Z1 Axio Observer, which was equipped with an incubator to maintain the samples at 37 °C and in 5% CO_2. A Zeiss illuminator HXP 120 V with an intensity of 80 % was used for imaging EYFP, mCherry and ECFP. The detection wavelength (band-pass filter set) for collecting emission light of EYFP was 520-550nm, for mCherry was 612-682nm and for ECFP was 460-500nm. Images were taken at a frequency of 1 image/10min. The detailed information for imaging was listed (Table 8-2). Channel 1, Channel 2 and Channel 3 correspond to EYFP, mCherry and ECFP.

Table 8-2 Acquisition information for imaging

Microscope	Axio Observer.Z1 /7			
Objective	LD Plan-Neofluar 20x/0.4 Korr Ph2 M27			
Image Size (Pixel)	45963 x 25819			
Scaling (per Pixel)	0.661 µm x 0.661 µm			
Image Size (Scaled)	30.37 mm x 17.06 mm			
Bit Depth	14 Bit			
ROI Center Offset	X: 0 µm, Y: 0 µm			
Channels	4			
	Channel 1	Channel 2	Channel 3	Channel 4
Filter	46 HE	64 HE	47 HE	Analy. DIC Trans. light
Beam Splitter	FT 515 HE	FT 605 HE	FT 455 HE	
Filter Ex. Wavelength	488-513	575-600	424-448	
Filter Em. Wavelength	520-550	612-682	460-500	

Contrast Method	Fluorescence	Fluorescence	Fluorescence	Phase
Light Source	HXP 120 V	HXP 120 V	HXP 120 V	TL LED
Light Source Intensity	79.16%	79.16%	79.16%	5 Volt
Dye Name	EYFP	mCherry	ECFP	
Excitation Wavelength	513	587	433	
Emission Wavelength	527	610	475	
Imaging Device	Axiocam 506	Axiocam 506	Axiocam 506	Axiocam 506
Camera Adapter	1x Camera Adapter	1x Camera Adapter	1x Camera Adapter	1x Camera Adapter
Exposure Time	1.5s	400 ms	180 ms	8 ms
Depth of Focus	6.70 µm	7.63 µm	5.94 µm	6.88 µm
Binning Mode	3,3	3,3	3,3	3,3

8.11 Imaging analysis and quantification

The live cell imaging data were quantified by customized pipeline in CellProfiler (Version 3.0.0)[146,147] and Fiji (Version 1.0)[178].

For EYFP-Smad2 signal, although the ratio of nuclear and cytoplasmic intensity was used as the output in some publications[90,179], we noticed that the fluorescence signal in the cytoplasm was not as robust as it in the nucleus for growing, moving cells in our work. Unexpected peaks or fluctuations of cytoplasmic EYFP-Smad2 fluorescence occurred frequently, which were caused by cell shape changing or noise from neighbors. However, the nuclear EYFP-Smad2 fluorescence was much less affected by cell shape changing or noise from neighbors. Besides, cytoplasm segmentation was less precise than nucleus segmentation. The intensity in the cytoplasm was not uniformly distributed thus the imprecise cytoplasm segmentation added noise in quantification. Therefore, the nuclear EYFP-Smad2 dynamics was employed. For each cell, the average nuclear YFP signal of the two frames before TGFβ stimulation was used as the basal nuclear YFP. The nuclear YFP signal in other frames were normalized by the basal signal[179].

The CellProfiler pipeline utilized the H2B-ECFP fluorescence channel to segment the nuclei. The nuclear Smad2 and hGeminin(1/110) level are quantified

as the mean YFP and mCherry signal in the nuclear region, respectively. The success rate of automatic long-term cell tracking by CellProfiler was low, mostly due to the cell overlapping and division. We were interested in the Smad2 dynamics during the 3h TGFβ treatment. Thus we decided to quantitatively analyze the nuclear EYFP-Smad2 signal of cells during the 3h TGFβ treatment by CellProfiler. And then the information of cell lineages was obtained by manually long-term cell tracking using Fiji. We tracked each cells manually and recorded the frames that the mCherry signal appeared (G1/S transition) or disappeared (cell division). Cells were excluded if they moved out of the FOV (field of view) before G1/S transition, or they were in G1 at the end of the movie but the G1 durations were less than 40h.

Chapter 9

Appendix

9.1 The values of parameters p_G and p_M in PBM

The values of the parameters p_G and p_M in the PBM (In Chapter 3: 3.1.3 Table 3-3) were listed.

Table 9-1 Values of parameters p_G

i	$p_G(i)$	i	$p_G(i)$	i	$p_G(i)$	i	$p_G(i)$	i	$p_G(i)$	i	$p_G(i)$
1	2.84E-122	93	7.30E-05	185	7.65E-03	277	1.04E-02	369	1.05E-03	461	2.18E-05
2	1.54E-92	94	8.36E-05	186	7.75E-03	278	1.03E-02	370	1.01E-03	462	2.09E-05
3	4.31E-77	95	9.54E-05	187	7.84E-03	279	1.02E-02	371	9.68E-04	463	2.00E-05
4	5.22E-67	96	1.08E-04	188	7.94E-03	280	1.01E-02	372	9.31E-04	464	1.92E-05
5	1.09E-59	97	1.23E-04	189	8.03E-03	281	1.01E-02	373	8.95E-04	465	1.84E-05
6	4.94E-54	98	1.39E-04	190	8.13E-03	282	9.98E-03	374	8.60E-04	466	1.76E-05
7	1.77E-49	99	1.56E-04	191	8.22E-03	283	9.89E-03	375	8.27E-04	467	1.68E-05
8	1.06E-45	100	1.75E-04	192	8.31E-03	284	9.80E-03	376	7.94E-04	468	1.61E-05
9	1.68E-42	101	1.96E-04	193	8.40E-03	285	9.70E-03	377	7.63E-04	469	1.54E-05
10	9.64E-40	102	2.18E-04	194	8.49E-03	286	9.61E-03	378	7.33E-04	470	1.48E-05
11	2.49E-37	103	2.42E-04	195	8.58E-03	287	9.51E-03	379	7.04E-04	471	1.42E-05
12	3.37E-35	104	2.68E-04	196	8.67E-03	288	9.41E-03	380	6.77E-04	472	1.36E-05
13	2.68E-33	105	2.97E-04	197	8.75E-03	289	9.30E-03	381	6.50E-04	473	1.30E-05
14	1.37E-31	106	3.27E-04	198	8.84E-03	290	9.20E-03	382	6.24E-04	474	1.24E-05
15	4.84E-30	107	3.59E-04	199	8.92E-03	291	9.09E-03	383	5.99E-04	475	1.19E-05
16	1.24E-28	108	3.94E-04	200	9.00E-03	292	8.98E-03	384	5.75E-04	476	1.14E-05
17	2.41E-27	109	4.31E-04	201	9.08E-03	293	8.86E-03	385	5.52E-04	477	1.09E-05
18	3.68E-26	110	4.70E-04	202	9.16E-03	294	8.75E-03	386	5.30E-04	478	1.05E-05
19	4.56E-25	111	5.11E-04	203	9.24E-03	295	8.63E-03	387	5.09E-04	479	1.00E-05
20	4.69E-24	112	5.55E-04	204	9.32E-03	296	8.51E-03	388	4.88E-04	480	9.60E-06
21	4.09E-23	113	6.01E-04	205	9.40E-03	297	8.39E-03	389	4.69E-04	481	9.19E-06
22	3.08E-22	114	6.50E-04	206	9.47E-03	298	8.26E-03	390	4.50E-04	482	8.80E-06
23	2.03E-21	115	7.01E-04	207	9.55E-03	299	8.14E-03	391	4.32E-04	483	8.43E-06
24	1.19E-20	116	7.55E-04	208	9.62E-03	300	8.01E-03	392	4.14E-04	484	8.07E-06
25	6.23E-20	117	8.11E-04	209	9.69E-03	301	7.88E-03	393	3.97E-04	485	7.73E-06
26	2.97E-19	118	8.70E-04	210	9.76E-03	302	7.75E-03	394	3.81E-04	486	7.40E-06
27	1.29E-18	119	9.31E-04	211	9.83E-03	303	7.62E-03	395	3.66E-04	487	7.09E-06
28	5.17E-18	120	9.94E-04	212	9.90E-03	304	7.49E-03	396	3.51E-04	488	6.79E-06
29	1.92E-17	121	1.06E-03	213	9.96E-03	305	7.36E-03	397	3.36E-04	489	6.50E-06
30	6.68E-17	122	1.13E-03	214	1.00E-02	306	7.22E-03	398	3.23E-04	490	6.23E-06
31	2.17E-16	123	1.20E-03	215	1.01E-02	307	7.09E-03	399	3.10E-04	491	5.96E-06
32	6.67E-16	124	1.27E-03	216	1.02E-02	308	6.95E-03	400	2.97E-04	492	5.71E-06
33	1.94E-15	125	1.35E-03	217	1.02E-02	309	6.82E-03	401	2.85E-04	493	5.47E-06
34	5.35E-15	126	1.43E-03	218	1.03E-02	310	6.68E-03	402	2.73E-04	494	5.24E-06
35	1.41E-14	127	1.51E-03	219	1.03E-02	311	6.54E-03	403	2.62E-04	495	5.02E-06
36	3.55E-14	128	1.59E-03	220	1.04E-02	312	6.41E-03	404	2.51E-04	496	4.81E-06

i	$p_G(i)$	i	$p_G(i)$	i	$p_G(i)$	i	$p_G(i)$	i	$p_G(i)$	i	$p_G(i)$
37	8.58E-14	129	1.67E-03	221	1.04E-02	313	6.27E-03	405	2.41E-04	497	4.60E-06
38	1.99E-13	130	1.76E-03	222	1.05E-02	314	6.13E-03	406	2.31E-04	498	4.41E-06
39	4.47E-13	131	1.85E-03	223	1.05E-02	315	6.00E-03	407	2.21E-04	499	4.22E-06
40	9.70E-13	132	1.94E-03	224	1.06E-02	316	5.86E-03	408	2.12E-04	500	4.04E-06
41	2.04E-12	133	2.03E-03	225	1.06E-02	317	5.72E-03	409	2.03E-04	501	3.87E-06
42	4.14E-12	134	2.12E-03	226	1.07E-02	318	5.59E-03	410	1.95E-04	502	3.71E-06
43	8.20E-12	135	2.22E-03	227	1.07E-02	319	5.46E-03	411	1.87E-04	503	3.55E-06
44	1.58E-11	136	2.32E-03	228	1.08E-02	320	5.32E-03	412	1.79E-04	504	3.40E-06
45	2.97E-11	137	2.42E-03	229	1.08E-02	321	5.19E-03	413	1.72E-04	505	3.26E-06
46	5.45E-11	138	2.52E-03	230	1.09E-02	322	5.06E-03	414	1.64E-04	506	3.12E-06
47	9.76E-11	139	2.62E-03	231	1.09E-02	323	4.93E-03	415	1.58E-04	507	2.99E-06
48	1.71E-10	140	2.72E-03	232	1.09E-02	324	4.80E-03	416	1.51E-04	508	2.86E-06
49	2.94E-10	141	2.83E-03	233	1.10E-02	325	4.67E-03	417	1.45E-04	509	2.74E-06
50	4.96E-10	142	2.93E-03	234	1.10E-02	326	4.55E-03	418	1.39E-04	510	2.62E-06
51	8.20E-10	143	3.04E-03	235	1.10E-02	327	4.43E-03	419	1.33E-04	511	2.51E-06
52	1.33E-09	144	3.15E-03	236	1.11E-02	328	4.30E-03	420	1.27E-04	512	2.41E-06
53	2.13E-09	145	3.26E-03	237	1.11E-02	329	4.18E-03	421	1.22E-04	513	2.31E-06
54	3.35E-09	146	3.37E-03	238	1.11E-02	330	4.06E-03	422	1.17E-04	514	2.21E-06
55	5.18E-09	147	3.48E-03	239	1.11E-02	331	3.95E-03	423	1.12E-04	515	2.12E-06
56	7.91E-09	148	3.59E-03	240	1.12E-02	332	3.83E-03	424	1.07E-04	516	2.03E-06
57	1.19E-08	149	3.70E-03	241	1.12E-02	333	3.72E-03	425	1.03E-04	517	1.94E-06
58	1.77E-08	150	3.81E-03	242	1.12E-02	334	3.61E-03	426	9.86E-05	518	1.86E-06
59	2.59E-08	151	3.92E-03	243	1.12E-02	335	3.50E-03	427	9.44E-05	519	1.78E-06
60	3.75E-08	152	4.04E-03	244	1.12E-02	336	3.39E-03	428	9.05E-05	520	1.71E-06
61	5.36E-08	153	4.15E-03	245	1.13E-02	337	3.29E-03	429	8.67E-05	521	1.63E-06
62	7.57E-08	154	4.26E-03	246	1.13E-02	338	3.19E-03	430	8.30E-05	522	1.56E-06
63	1.06E-07	155	4.38E-03	247	1.13E-02	339	3.09E-03	431	7.95E-05	523	1.50E-06
64	1.47E-07	156	4.49E-03	248	1.13E-02	340	2.99E-03	432	7.62E-05	524	1.44E-06
65	2.01E-07	157	4.61E-03	249	1.13E-02	341	2.89E-03	433	7.30E-05	525	1.38E-06
66	2.73E-07	158	4.72E-03	250	1.13E-02	342	2.80E-03	434	6.99E-05	526	1.32E-06
67	3.67E-07	159	4.84E-03	251	1.13E-02	343	2.71E-03	435	6.70E-05	527	1.26E-06
68	4.89E-07	160	4.95E-03	252	1.13E-02	344	2.62E-03	436	6.42E-05	528	1.21E-06
69	6.46E-07	161	5.06E-03	253	1.13E-02	345	2.53E-03	437	6.15E-05	529	1.16E-06
70	8.46E-07	162	5.18E-03	254	1.13E-02	346	2.45E-03	438	5.89E-05	530	1.11E-06
71	1.10E-06	163	5.29E-03	255	1.13E-02	347	2.36E-03	439	5.64E-05	531	1.06E-06
72	1.42E-06	164	5.40E-03	256	1.13E-02	348	2.28E-03	440	5.40E-05	532	1.02E-06
73	1.81E-06	165	5.52E-03	257	1.12E-02	349	2.20E-03	441	5.18E-05	533	9.75E-07
74	2.31E-06	166	5.63E-03	258	1.12E-02	350	2.13E-03	442	4.96E-05	534	9.34E-07
75	2.91E-06	167	5.74E-03	259	1.12E-02	351	2.05E-03	443	4.75E-05	535	8.95E-07
76	3.65E-06	168	5.85E-03	260	1.12E-02	352	1.98E-03	444	4.55E-05	536	8.57E-07
77	4.54E-06	169	5.96E-03	261	1.12E-02	353	1.91E-03	445	4.36E-05	537	8.21E-07
78	5.62E-06	170	6.08E-03	262	1.11E-02	354	1.84E-03	446	4.17E-05	538	7.86E-07

i	$p_G(i)$	i	$p_G(i)$	i	$p_G(i)$	i	$p_G(i)$	i	$p_G(i)$	i	$p_G(i)$
79	6.92E-06	171	6.19E-03	263	1.11E-02	355	1.78E-03	447	4.00E-05	539	7.53E-07
80	8.46E-06	172	6.30E-03	264	1.11E-02	356	1.71E-03	448	3.83E-05	540	7.22E-07
81	1.03E-05	173	6.40E-03	265	1.10E-02	357	1.65E-03	449	3.67E-05	541	6.91E-07
82	1.24E-05	174	6.51E-03	266	1.10E-02	358	1.59E-03	450	3.51E-05	542	6.62E-07
83	1.50E-05	175	6.62E-03	267	1.09E-02	359	1.53E-03	451	3.36E-05	543	6.35E-07
84	1.79E-05	176	6.73E-03	268	1.09E-02	360	1.48E-03	452	3.22E-05	544	6.08E-07
85	2.13E-05	177	6.83E-03	269	1.09E-02	361	1.42E-03	453	3.08E-05	545	5.82E-07
86	2.52E-05	178	6.94E-03	270	1.08E-02	362	1.37E-03	454	2.95E-05	546	5.58E-07
87	2.98E-05	179	7.04E-03	271	1.08E-02	363	1.32E-03	455	2.83E-05	547	5.35E-07
88	3.49E-05	180	7.14E-03	272	1.07E-02	364	1.27E-03	456	2.71E-05	548	5.12E-07
89	4.08E-05	181	7.25E-03	273	1.06E-02	365	1.22E-03	457	2.60E-05	549	4.91E-07
90	4.75E-05	182	7.35E-03	274	1.06E-02	366	1.18E-03	458	2.49E-05	550	4.70E-07
91	5.51E-05	183	7.45E-03	275	1.05E-02	367	1.13E-03	459	2.38E-05		
92	6.35E-05	184	7.55E-03	276	1.04E-02	368	1.09E-03	460	2.28E-05		

Table 9-2 Values of parameters p_M

j	$p_M(j)$	j	$p_M(j)$	j	$p_M(j)$	j	$p_M(j)$	j	$p_M(j)$	j	$p_M(j)$
1	6.70E-315	108	7.34E-13	215	3.91E-04	322	8.39E-03	429	7.45E-03	536	1.14E-04
2	9.10E-244	109	1.10E-12	216	4.18E-04	323	8.47E-03	430	7.31E-03	537	1.08E-04
3	2.25E-206	110	1.65E-12	217	4.45E-04	324	8.55E-03	431	7.17E-03	538	1.03E-04
4	9.64E-182	111	2.44E-12	218	4.74E-04	325	8.63E-03	432	7.02E-03	539	9.81E-05
5	1.00E-163	112	3.59E-12	219	5.04E-04	326	8.71E-03	433	6.88E-03	540	9.33E-05
6	1.04E-149	113	5.25E-12	220	5.35E-04	327	8.78E-03	434	6.73E-03	541	8.88E-05
7	2.33E-138	114	7.63E-12	221	5.68E-04	328	8.86E-03	435	6.58E-03	542	8.45E-05
8	6.83E-129	115	1.10E-11	222	6.02E-04	329	8.94E-03	436	6.44E-03	543	8.04E-05
9	7.97E-121	116	1.57E-11	223	6.37E-04	330	9.01E-03	437	6.29E-03	544	7.65E-05
10	7.82E-114	117	2.24E-11	224	6.74E-04	331	9.08E-03	438	6.14E-03	545	7.28E-05
11	1.08E-107	118	3.17E-11	225	7.12E-04	332	9.16E-03	439	6.00E-03	546	6.92E-05
12	3.09E-102	119	4.46E-11	226	7.51E-04	333	9.23E-03	440	5.85E-03	547	6.58E-05
13	2.40E-97	120	6.23E-11	227	7.92E-04	334	9.30E-03	441	5.70E-03	548	6.26E-05
14	6.28E-93	121	8.65E-11	228	8.34E-04	335	9.37E-03	442	5.56E-03	549	5.96E-05
15	6.56E-89	122	1.20E-10	229	8.78E-04	336	9.44E-03	443	5.42E-03	550	5.66E-05
16	3.12E-85	123	1.64E-10	230	9.23E-04	337	9.51E-03	444	5.27E-03	551	5.39E-05
17	7.51E-82	124	2.24E-10	231	9.69E-04	338	9.58E-03	445	5.13E-03	552	5.12E-05
18	9.99E-79	125	3.05E-10	232	1.02E-03	339	9.65E-03	446	4.99E-03	553	4.87E-05
19	7.88E-76	126	4.12E-10	233	1.07E-03	340	9.71E-03	447	4.85E-03	554	4.63E-05
20	3.92E-73	127	5.54E-10	234	1.12E-03	341	9.78E-03	448	4.71E-03	555	4.40E-05
21	1.29E-70	128	7.41E-10	235	1.17E-03	342	9.84E-03	449	4.58E-03	556	4.18E-05
22	2.95E-68	129	9.86E-10	236	1.22E-03	343	9.90E-03	450	4.44E-03	557	3.98E-05
23	4.85E-66	130	1.31E-09	237	1.28E-03	344	9.96E-03	451	4.31E-03	558	3.78E-05

j	$p_M(j)$	j	$p_M(j)$	j	$p_M(j)$	j	$p_M(j)$	j	$p_M(j)$	j	$p_M(j)$
24	5.89E-64	131	1.72E-09	238	1.33E-03	345	1.00E-02	452	4.18E-03	559	3.59E-05
25	5.47E-62	132	2.26E-09	239	1.39E-03	346	1.01E-02	453	4.05E-03	560	3.42E-05
26	3.96E-60	133	2.95E-09	240	1.45E-03	347	1.01E-02	454	3.93E-03	561	3.25E-05
27	2.29E-58	134	3.84E-09	241	1.51E-03	348	1.02E-02	455	3.80E-03	562	3.09E-05
28	1.07E-56	135	4.97E-09	242	1.57E-03	349	1.03E-02	456	3.68E-03	563	2.93E-05
29	4.16E-55	136	6.40E-09	243	1.64E-03	350	1.03E-02	457	3.56E-03	564	2.79E-05
30	1.35E-53	137	8.22E-09	244	1.70E-03	351	1.04E-02	458	3.44E-03	565	2.65E-05
31	3.74E-52	138	1.05E-08	245	1.77E-03	352	1.04E-02	459	3.33E-03	566	2.52E-05
32	8.88E-51	139	1.34E-08	246	1.83E-03	353	1.05E-02	460	3.22E-03	567	2.39E-05
33	1.83E-49	140	1.69E-08	247	1.90E-03	354	1.05E-02	461	3.11E-03	568	2.27E-05
34	3.32E-48	141	2.14E-08	248	1.97E-03	355	1.06E-02	462	3.00E-03	569	2.16E-05
35	5.32E-47	142	2.69E-08	249	2.04E-03	356	1.06E-02	463	2.90E-03	570	2.05E-05
36	7.60E-46	143	3.37E-08	250	2.11E-03	357	1.07E-02	464	2.79E-03	571	1.95E-05
37	9.77E-45	144	4.21E-08	251	2.19E-03	358	1.07E-02	465	2.69E-03	572	1.85E-05
38	1.13E-43	145	5.24E-08	252	2.26E-03	359	1.07E-02	466	2.60E-03	573	1.76E-05
39	1.20E-42	146	6.49E-08	253	2.33E-03	360	1.08E-02	467	2.50E-03	574	1.67E-05
40	1.16E-41	147	8.02E-08	254	2.41E-03	361	1.08E-02	468	2.41E-03	575	1.58E-05
41	1.03E-40	148	9.86E-08	255	2.49E-03	362	1.09E-02	469	2.32E-03	576	1.50E-05
42	8.49E-40	149	1.21E-07	256	2.56E-03	363	1.09E-02	470	2.24E-03	577	1.43E-05
43	6.48E-39	150	1.48E-07	257	2.64E-03	364	1.09E-02	471	2.15E-03	578	1.36E-05
44	4.60E-38	151	1.80E-07	258	2.72E-03	365	1.10E-02	472	2.07E-03	579	1.29E-05
45	3.06E-37	152	2.19E-07	259	2.80E-03	366	1.10E-02	473	1.99E-03	580	1.22E-05
46	1.91E-36	153	2.65E-07	260	2.89E-03	367	1.10E-02	474	1.91E-03	581	1.16E-05
47	1.12E-35	154	3.19E-07	261	2.97E-03	368	1.10E-02	475	1.84E-03	582	1.10E-05
48	6.21E-35	155	3.84E-07	262	3.05E-03	369	1.11E-02	476	1.77E-03	583	1.05E-05
49	3.26E-34	156	4.61E-07	263	3.13E-03	370	1.11E-02	477	1.70E-03	584	9.94E-06
50	1.63E-33	157	5.51E-07	264	3.22E-03	371	1.11E-02	478	1.63E-03	585	9.44E-06
51	7.71E-33	158	6.57E-07	265	3.30E-03	372	1.11E-02	479	1.56E-03	586	8.96E-06
52	3.49E-32	159	7.81E-07	266	3.39E-03	373	1.11E-02	480	1.50E-03	587	8.50E-06
53	1.51E-31	160	9.26E-07	267	3.48E-03	374	1.12E-02	481	1.44E-03	588	8.07E-06
54	6.24E-31	161	1.09E-06	268	3.56E-03	375	1.12E-02	482	1.38E-03	589	7.66E-06
55	2.48E-30	162	1.29E-06	269	3.65E-03	376	1.12E-02	483	1.33E-03	590	7.27E-06
56	9.47E-30	163	1.52E-06	270	3.74E-03	377	1.12E-02	484	1.27E-03	591	6.90E-06
57	3.48E-29	164	1.78E-06	271	3.83E-03	378	1.12E-02	485	1.22E-03	592	6.55E-06
58	1.23E-28	165	2.08E-06	272	3.92E-03	379	1.12E-02	486	1.17E-03	593	6.22E-06
59	4.22E-28	166	2.43E-06	273	4.01E-03	380	1.12E-02	487	1.12E-03	594	5.90E-06
60	1.40E-27	167	2.82E-06	274	4.10E-03	381	1.12E-02	488	1.07E-03	595	5.60E-06
61	4.48E-27	168	3.28E-06	275	4.19E-03	382	1.12E-02	489	1.03E-03	596	5.31E-06
62	1.39E-26	169	3.79E-06	276	4.28E-03	383	1.12E-02	490	9.83E-04	597	5.04E-06
63	4.20E-26	170	4.38E-06	277	4.37E-03	384	1.12E-02	491	9.41E-04	598	4.78E-06
64	1.23E-25	171	5.05E-06	278	4.46E-03	385	1.11E-02	492	9.00E-04	599	4.54E-06
65	3.51E-25	172	5.80E-06	279	4.55E-03	386	1.11E-02	493	8.62E-04	600	4.31E-06

j	$p_M(j)$	j	$p_M(j)$	j	$p_M(j)$	j	$p_M(j)$	j	$p_M(j)$	j	$p_M(j)$
66	9.76E-25	173	6.66E-06	280	4.64E-03	387	1.11E-02	494	8.24E-04	601	4.08E-06
67	2.64E-24	174	7.62E-06	281	4.74E-03	388	1.11E-02	495	7.89E-04	602	3.87E-06
68	6.98E-24	175	8.70E-06	282	4.83E-03	389	1.11E-02	496	7.54E-04	603	3.68E-06
69	1.80E-23	176	9.91E-06	283	4.92E-03	390	1.10E-02	497	7.21E-04	604	3.49E-06
70	4.54E-23	177	1.13E-05	284	5.01E-03	391	1.10E-02	498	6.90E-04	605	3.31E-06
71	1.12E-22	178	1.28E-05	285	5.11E-03	392	1.10E-02	499	6.59E-04	606	3.14E-06
72	2.70E-22	179	1.45E-05	286	5.20E-03	393	1.09E-02	500	6.30E-04	607	2.98E-06
73	6.38E-22	180	1.63E-05	287	5.29E-03	394	1.09E-02	501	6.02E-04	608	2.82E-06
74	1.48E-21	181	1.84E-05	288	5.38E-03	395	1.08E-02	502	5.75E-04	609	2.68E-06
75	3.36E-21	182	2.07E-05	289	5.48E-03	396	1.08E-02	503	5.50E-04	610	2.54E-06
76	7.48E-21	183	2.33E-05	290	5.57E-03	397	1.07E-02	504	5.25E-04	611	2.41E-06
77	1.64E-20	184	2.61E-05	291	5.66E-03	398	1.07E-02	505	5.01E-04	612	2.28E-06
78	3.52E-20	185	2.92E-05	292	5.76E-03	399	1.06E-02	506	4.79E-04	613	2.17E-06
79	7.45E-20	186	3.26E-05	293	5.85E-03	400	1.06E-02	507	4.57E-04	614	2.05E-06
80	1.55E-19	187	3.63E-05	294	5.94E-03	401	1.05E-02	508	4.36E-04	615	1.95E-06
81	3.17E-19	188	4.04E-05	295	6.03E-03	402	1.04E-02	509	4.16E-04	616	1.85E-06
82	6.39E-19	189	4.48E-05	296	6.12E-03	403	1.04E-02	510	3.97E-04	617	1.75E-06
83	1.27E-18	190	4.97E-05	297	6.22E-03	404	1.03E-02	511	3.79E-04	618	1.66E-06
84	2.48E-18	191	5.50E-05	298	6.31E-03	405	1.02E-02	512	3.62E-04	619	1.57E-06
85	4.78E-18	192	6.07E-05	299	6.40E-03	406	1.01E-02	513	3.45E-04	620	1.49E-06
86	9.10E-18	193	6.69E-05	300	6.49E-03	407	1.01E-02	514	3.29E-04	621	1.42E-06
87	1.71E-17	194	7.37E-05	301	6.58E-03	408	9.97E-03	515	3.14E-04	622	1.34E-06
88	3.16E-17	195	8.10E-05	302	6.67E-03	409	9.88E-03	516	3.00E-04	623	1.27E-06
89	5.78E-17	196	8.88E-05	303	6.76E-03	410	9.78E-03	517	2.86E-04	624	1.21E-06
90	1.04E-16	197	9.73E-05	304	6.85E-03	411	9.68E-03	518	2.72E-04	625	1.14E-06
91	1.86E-16	198	1.06E-04	305	6.94E-03	412	9.58E-03	519	2.60E-04	626	1.08E-06
92	3.29E-16	199	1.16E-04	306	7.03E-03	413	9.48E-03	520	2.48E-04	627	1.03E-06
93	5.73E-16	200	1.27E-04	307	7.12E-03	414	9.37E-03	521	2.36E-04	628	9.74E-07
94	9.88E-16	201	1.38E-04	308	7.21E-03	415	9.26E-03	522	2.25E-04	629	9.24E-07
95	1.69E-15	202	1.50E-04	309	7.29E-03	416	9.15E-03	523	2.14E-04	630	8.76E-07
96	2.85E-15	203	1.63E-04	310	7.38E-03	417	9.03E-03	524	2.04E-04	631	8.30E-07
97	4.76E-15	204	1.76E-04	311	7.47E-03	418	8.91E-03	525	1.95E-04	632	7.87E-07
98	7.87E-15	205	1.91E-04	312	7.56E-03	419	8.79E-03	526	1.85E-04	633	7.46E-07
99	1.29E-14	206	2.06E-04	313	7.64E-03	420	8.67E-03	527	1.77E-04	634	7.07E-07
100	2.09E-14	207	2.23E-04	314	7.73E-03	421	8.54E-03	528	1.68E-04	635	6.70E-07
101	3.36E-14	208	2.40E-04	315	7.81E-03	422	8.41E-03	529	1.60E-04	636	6.35E-07
102	5.36E-14	209	2.58E-04	316	7.90E-03	423	8.28E-03	530	1.53E-04	637	6.02E-07
103	8.46E-14	210	2.78E-04	317	7.98E-03	424	8.15E-03	531	1.45E-04		
104	1.33E-13	211	2.98E-04	318	8.06E-03	425	8.01E-03	532	1.38E-04		
105	2.06E-13	212	3.20E-04	319	8.15E-03	426	7.87E-03	533	1.32E-04		
106	3.17E-13	213	3.42E-04	320	8.23E-03	427	7.74E-03	534	1.25E-04		
107	4.84E-13	214	3.66E-04	321	8.31E-03	428	7.59E-03	535	1.19E-04		

9.2 Parameter values for arrested cells in PBM Model 3 and Model4

In Chapter 3: 3.2.2.2 , we simulated PBM model 3 and Model 4. We did three pairs of simulations for comparison ((1) $G_{Sen} = M_{Sen}$; (2) $G_{Sen} < M_{Sen}$; (3) $G_{Sen} > M_{Sen}$). For each pair of simulation, the values of parameters were listed and initial G1 phase durations were simulated (Table 9-3, Table 9-4 and Table 9-5).

Table 9-3 Values of parameters in Model 3 and Model4 for simulation (1)

Graded response						
	1st G1 (Stimulated at G1 phase)			2nd G1 (Stimulated at S/G2/M phase)		
	p'_G	D'_{1stG} (h)	STD'_{1stG} (h)	p'_M	D'_{2ndG} (h)	STD'_{2ndG} (h)
1	0	8.1	1.8	0	8.1	1.8
2	1	12.3	1.8	1	12.3	1.8
3	1	16.5	1.8	1	16.5	1.8
4	1	20.7	1.8	1	20.7	1.8
5	1	24.9	1.8	1	24.9	1.8
6	1	29.1	1.8	1	29.1	1.8
7	1	33.2	1.8	1	33.2	1.8
8	1	37.4	1.8	1	37.4	1.8
9	1	41.6	1.8	1	41.6	1.8
10	1	45.8	1.8	1	45.8	1.8
11	1	50	1.8	1	50	1.8
All-or-none response						
	1st G1 (Stimulated at G1 phase)			2nd G1 (Stimulated at S/G2/M phase)		
	p'_G	D'_G (h)	STD'_G (h)	p'_M	D'_{2ndG} (h)	STD'_{2ndG} (h)
1	0	8.1	1.8	0	8.1	1.8
2	0.1	40.5	4.5	0.1	40.5	4.5
3	0.2	40.5	4.5	0.2	40.5	4.5
4	0.3	40.5	4.5	0.3	40.5	4.5
5	0.4	40.5	4.5	0.4	40.5	4.5
6	0.5	40.5	4.5	0.5	40.5	4.5
7	0.6	40.5	4.5	0.6	40.5	4.5
8	0.7	40.5	4.5	0.7	40.5	4.5
9	0.8	40.5	4.5	0.8	40.5	4.5
10	0.9	40.5	4.5	0.9	40.5	4.5
11	1	40.5	4.5	1	40.5	4.5

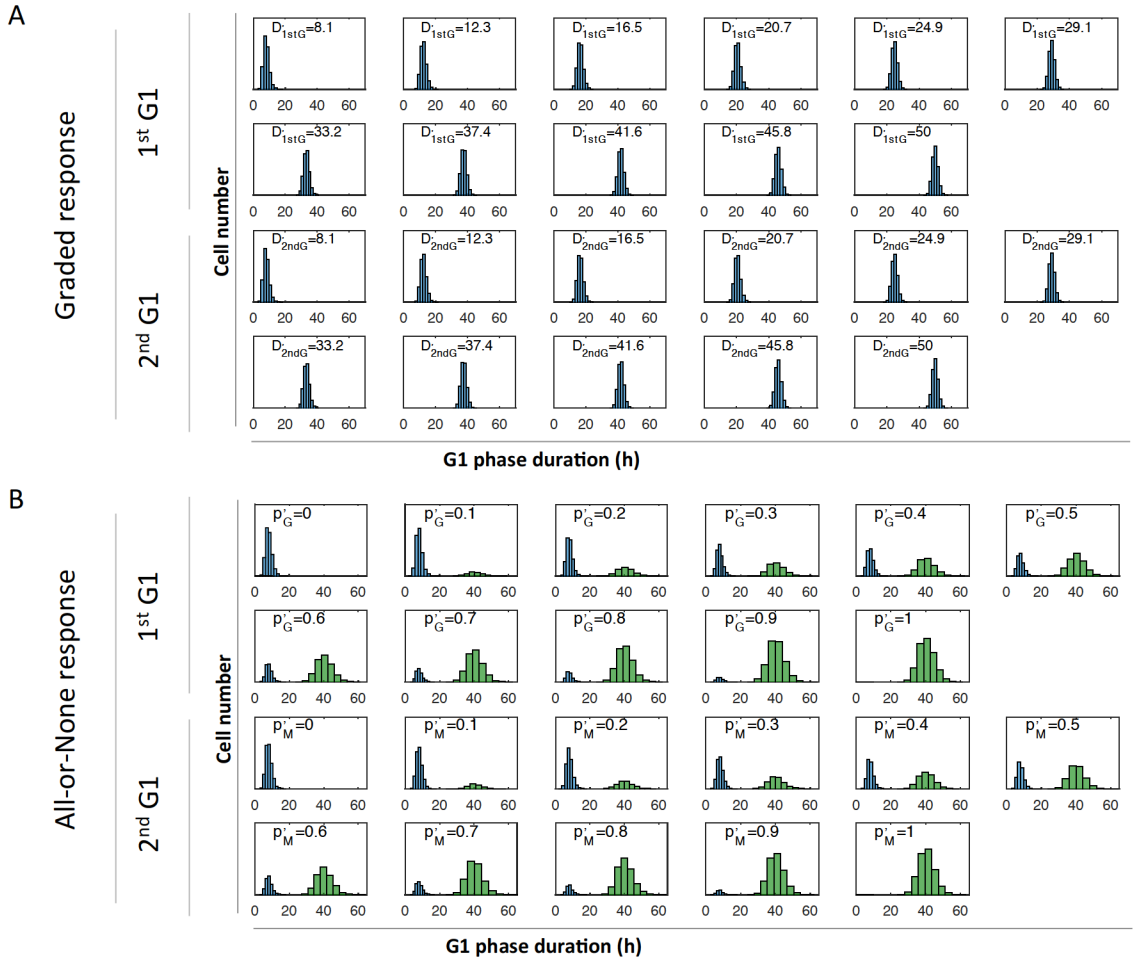


Figure 9.1 The G1 phase duration distributions of cells in the initial conditions of simulation (1). (A) Graded response; $D'_{1stG} = D'_{2ndG}$; the arrested G1 durations of the cells stimulated in G1 or S/G2/M phase were fitting to the same distribution; D'_{1stG}: the mean of arrested G1 phase duration for the cells arrested at 1st G1; D'_{2ndG}: the mean of arrested G1 phase duration for the cells arrested at 2nd G1. (B) All-or-none response; $p'_G = p'_M$; the cells stimulated in G1 or S/G2/M phase had the same probabilities to be arrested; p'_G: the probability of being arrested when stimulated in G1 phase; p'_M: the probability of being affected when stimulated in S/G2/M phase; Blue distributions: unaffected cells; Green distributions: arrested cells.

Table 9-4 Values of parameters in Model 3 and Model4 for simulation (2)

Graded response						
	1st G1 (Stimulated at G1 phase)			2nd G1 (Stimulated at S/G2/M phase)		
	p'_G	D'_{1stG} (h)	STD'_{1stG} (h)	p'_M	D'_{2ndG} (h)	STD'_{2ndG} (h)
1	0	8.1	1.8	0	8.1	1.8
2	1	12.3	1.8	1	13.3	1.8
3	1	16.5	1.8	1	18.5	1.8
4	1	20.7	1.8	1	23.7	1.8
5	1	24.9	1.8	1	28.9	1.8
6	1	29.1	1.8	1	34.1	1.8
7	1	33.2	1.8	1	39.2	1.8
8	1	37.4	1.8	1	44.4	1.8
9	1	41.6	1.8	1	49.6	1.8
10	1	45.8	1.8	1	54.8	1.8
11	1	50	1.8	1	60.0	1.8
All-or-none response						
	1st G1 (Stimulated at G1 phase)			2nd G1 (Stimulated at S/G2/M phase)		
	p'_G	D'_G (h)	STD'_G (h)	p'_M	D'_{2ndG} (h)	STD'_{2ndG} (h)
1	0	8.1	1.8	0	8.1	1.8
2	0.08	40.5	4.5	0.1	40.5	4.5
3	0.16	40.5	4.5	0.2	40.5	4.5
4	0.24	40.5	4.5	0.3	40.5	4.5
5	0.32	40.5	4.5	0.4	40.5	4.5
6	0.4	40.5	4.5	0.5	40.5	4.5
7	0.48	40.5	4.5	0.6	40.5	4.5
8	0.56	40.5	4.5	0.7	40.5	4.5
9	0.64	40.5	4.5	0.8	40.5	4.5
10	0.72	40.5	4.5	0.9	40.5	4.5
11	0.8	40.5	4.5	1	40.5	4.5

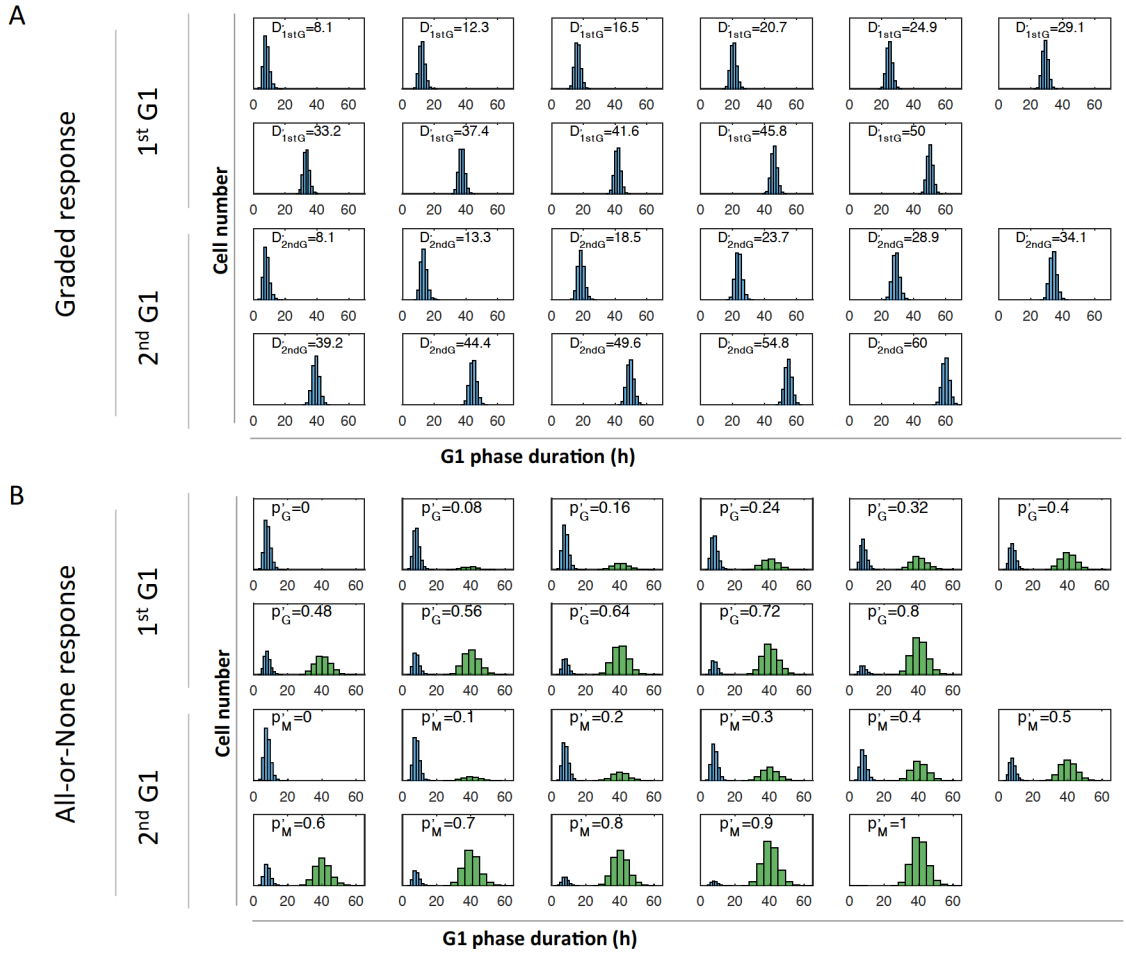


Figure 9.2 The G1 phase duration distributions of cells in the initial conditions of simulation (2). (A) (A) Graded response; $D'_{1stG} < D'_{2ndG}$; the average arrested G1 duration of the cells stimulated in G1 is shorter than cells stimulated at S/G2/M phase; D'_{1stG}: the mean of arrested G1 phase duration for the cells arrested at 1st G1; D'_{2ndG}: the mean of arrested G1 phase duration for the cells arrested at 2nd G1. (B) All-or-none response; $p'_G < p'_M$; the cells stimulated in G1 phase had the smaller probability to be arrested than in S/G2/M phase p'_G: the probability of being arrested when stimulated in G1 phase; p'_M: the probability of being affected when stimulated in S/G2/M phase; Blue distributions: unaffected cells; Green distributions: arrested cells.

Table 9-5 Values of parameters in Model 3 and Model4 for simulation (3)

Graded response						
	1st G1 (Stimulated at G1 phase)			2nd G1 (Stimulated at S/G2/M phase)		
	p'_G	D'_{1stG} (h)	STD'_{1stG} (h)	p'_M	D'_{2ndG} (h)	STD'_{2ndG} (h)
1	0	8.1	1.8	0	8.1	1.8
2	1	12.3	1.8	1	11.3	1.8
3	1	16.5	1.8	1	14.5	1.8
4	1	20.7	1.8	1	17.7	1.8
5	1	24.9	1.8	1	20.9	1.8
6	1	29.1	1.8	1	24.1	1.8
7	1	33.2	1.8	1	27.2	1.8
8	1	37.4	1.8	1	30.4	1.8
9	1	41.6	1.8	1	33.6	1.8
10	1	45.8	1.8	1	36.8	1.8
11	1	50	1.8	1	40.0	1.8
All-or-none response						
	1st G1 (Stimulated at G1 phase)			2nd G1 (Stimulated at S/G2/M phase)		
	p'_G	D'_G (h)	STD'_G (h)	p'_M	D'_{2ndG} (h)	STD'_{2ndG} (h)
1	0	8.1	1.8	0	8.1	1.8
2	0.1	40.5	4.5	0.08	40.5	4.5
3	0.2	40.5	4.5	0.16	40.5	4.5
4	0.3	40.5	4.5	0.24	40.5	4.5
5	0.4	40.5	4.5	0.32	40.5	4.5
6	0.5	40.5	4.5	0.4	40.5	4.5
7	0.6	40.5	4.5	0.48	40.5	4.5
8	0.7	40.5	4.5	0.56	40.5	4.5
9	0.8	40.5	4.5	0.64	40.5	4.5
10	0.9	40.5	4.5	0.72	40.5	4.5
11	1	40.5	4.5	0.8	40.5	4.5

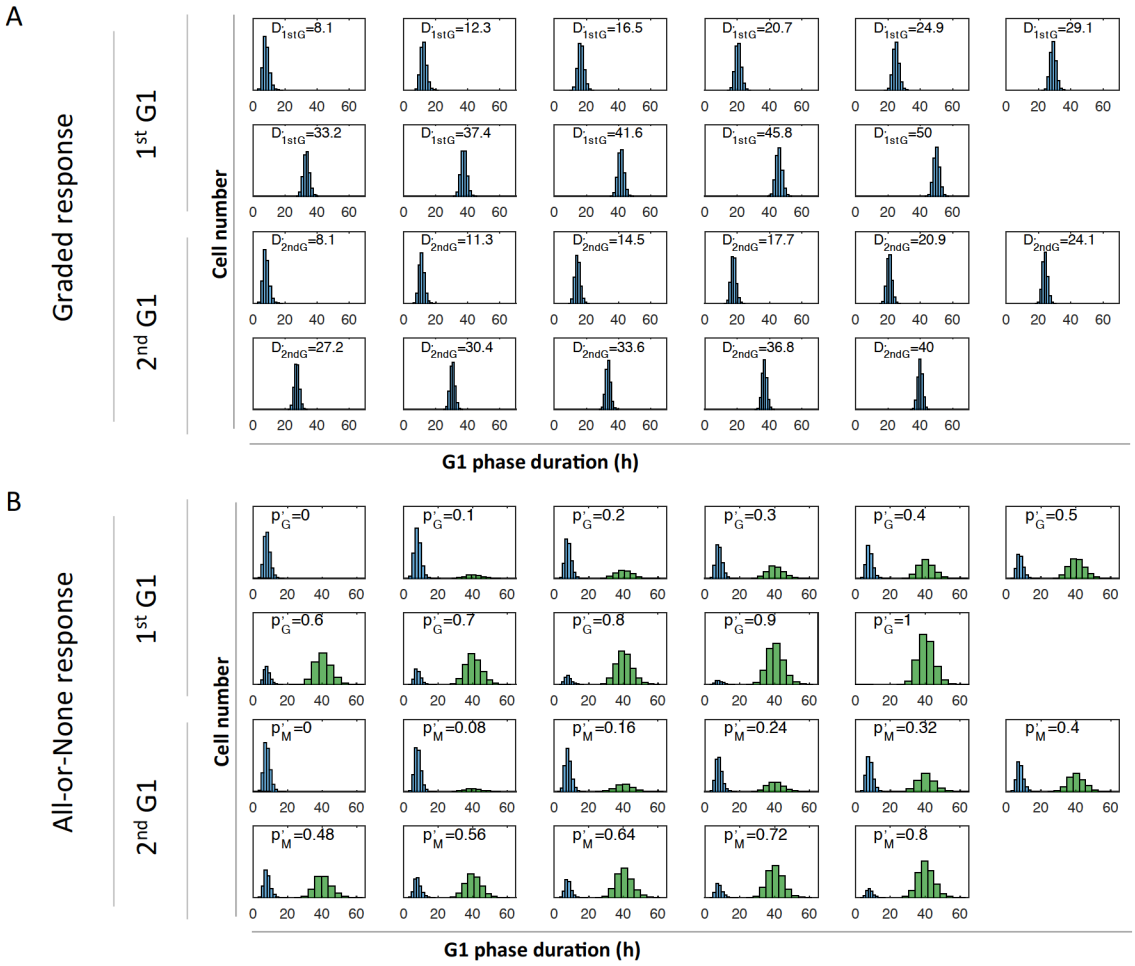


Figure 9.3 The G1 phase duration distributions of cells in the initial conditions of simulation (3). (A) Graded response; $D'_{1stG} > D'_{2ndG}$; the average arrested G1 duration of the cells stimulated in G1 is longer than cells stimulated at S/G2/M phase; D'_{1stG}: the mean of arrested G1 phase duration for the cells arrested at 1st G1; D'_{2ndG}: the mean of arrested G1 phase duration for the cells arrested at 2nd G1. (B) All-or-none response; $p'_G > p'_M$; the cells stimulated in G1 phase had the larger probability to be arrested than in S/G2/M phase; p'_G : the probability of being arrested when stimulated in G1 phase; p'_M: the probability of being affected when stimulated in S/G2/M phase; Blue distributions: unaffected cells; Green distributions: arrested cells.

9.3 Validation of method for data scaling

9.3.1 Method description

We found that the fraction of cells in G1 phase was not constant in the control group (untreated cells). There were some fluctuations (±10%) over time, caused by several ligand independent factors, e.g., medium change. Thus the curve describing the responses of experimental groups (R_{total}) actually included two parts: ligand dependent response (LDR) and ligand independent fluctuations (LIF). Since we focused on ligand dependent G1 arrest and only the ligand dependent response (LDR) was included in modeling, to compare model simulation with experimental data, we would like to exclude the ligand independent fluctuations (LIF) from the experimental data.

Mathematically, the LIF could be calculated by solving the equations (Equation 16 and 17). However, it was too complicate to solve them and calculated LIF of each experiment. The LIF was estimated by calculating the difference between the curve describing the G1 fraction over time of untreated cells in experiments ($C_{Experiment}$, including LIF) and that in model simulation (C_{Model}, quasi-steady state) (Equation 27, Figure 9.4).

$$LIF = C_{Experiment} - C_{Model} \tag{27}$$

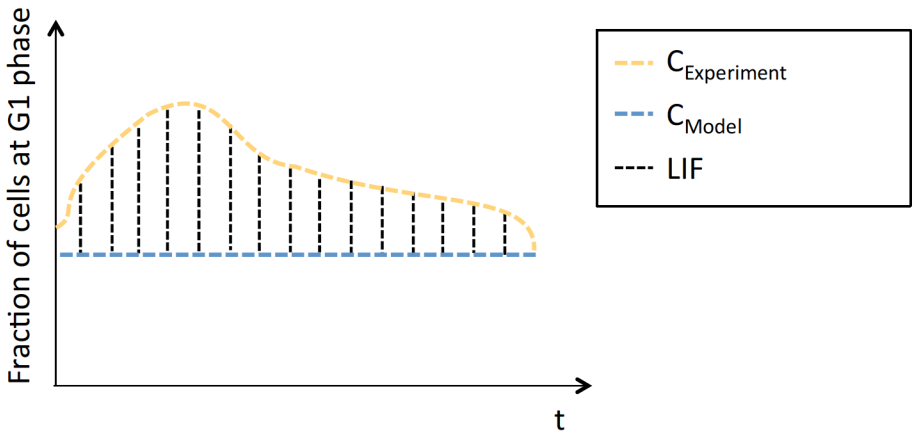


Figure 9.4 Schematic showing the method of estimating ligand independent fluctuations (LIF). The ligand independent fluctuations (LIF) were estimated by calculating the difference between the control data of experiments ($C_{Experiment}$) and model (C_{Model}).

After that, the LDR of all the experimental data could be estimated by subtracting LIF from R_{total} (Equation 28). The LDR of experimental data and model simulations were compared.

$$LDR \approx LDR' = R_{total} - LIF \quad (28)$$

9.3.2 Method validation

Next, while this method was difficult to test experimentally, we validated this method by model simulation. We simulated pseudodatas LDR and R_{total}, then calculated LIF and LDR' (Table 9-6). If the LDR' and LDR were similar, we could use the LDR' to represent LDR. Here, 30 randomly selected sets of simulations were shown (Table 9-7 and Figure 9.5).

Table 9-6 Equations and annotations of simulated data

	Data	Equations and Annotations
1	$\boldsymbol{C}$	$\boldsymbol{f(0,0,0,0)}$ Control data: The pseudodata for the curve describing cells without treatment was simulated by setting AP parameters as p1= 0, p2= 0, p3= 0 and p4= 0.
2	$\boldsymbol{C'}$	$\boldsymbol{f(Noise\_p1, Noise\_p2, Noise\_p3, Noise\_p4)}$ Control data with LIF: The pseudodata for the curve describing cells without treatment with LIF was simulated by setting AP parameters as p1= *Noise_p1*, p2= *Noise_p2*, p3= *Noise_p3* and p4= *Noise_p4* (Table 9-7).
3	$\boldsymbol{LDR}$	$\boldsymbol{f(p1,p2,p3,p4)}$ Data of ligand-treated group: The pseudodata for the curve describing cells upon 50pM TGFβ treatment was simulated by setting AP parameters as p1=0.2895, p2=0.1053, p3=0.4231 and p4=0.0096.
4	$\boldsymbol{R_{total}}$	$\boldsymbol{f(p1 + Noise\_p1, p2 + Noise\_p2, p3 + Noise\_p3, p4 + Noise\_p4)}$ Data of ligand-treated group with LIF: The pseudodata for the curve describing cells upon 50pM TGFβ treatment with LIF was simulated by setting AP parameters as p1=0.2895+Noise_p1, p2=0.1053+Noise_p2, p3=0.4231+Noise_p3 and p4=0.0096+Noise_p4 (Table 9-7).
5	$\boldsymbol{LIF}$	$\boldsymbol{C_{Experiment} - C_{Model}}$ Estimated LIF: The pseudodata for the difference between the curve representing steady state and
6	$\boldsymbol{LDR'}$	$\boldsymbol{R_{total} - LIF}$ Estimated LDR: The LDR was estimated by calculating the difference between $\boldsymbol{R_{total}}$ and $\boldsymbol{LIF}$.

The value of $Noise\_p1, Noise\_p2, Noise\_p3$ *and* $Noise\_p4$ were obtained by Latin Hypercube sampling[143] (30 sets were uniformly sampled in the 4-dimensional space during the range of [0, 0.2], Table 9-7).

Table 9-7 The noises for four AP parameters

	Noise_p1	Noise_p2	Noise_p3	Noise_p4
1	0.0404	0.1135	0.0143	0.0464
2	0.1517	0.0829	0.1166	0.0964
3	0.1670	0.1519	0.0599	0.0618
4	0.1633	0.0143	0.0739	0.1055
5	0.1096	0.0409	0.1306	0.1505
6	0.1334	0.0979	0.0085	0.1393
7	0.1749	0.0512	0.0057	0.0846
8	0.0846	0.0611	0.0479	0.1815
9	0.0596	0.1034	0.1006	0.0698
10	0.1330	0.0263	0.1081	0.0882
11	0.0612	0.1681	0.0350	0.1291
12	0.1239	0.1638	0.1773	0.0074
13	0.0108	0.0715	0.1620	0.0529
14	0.1965	0.0378	0.0696	0.0388
15	0.0341	0.1109	0.1961	0.0239
16	0.1173	0.1280	0.1377	0.1884
17	0.0760	0.0797	0.0286	0.0007
18	0.1460	0.1818	0.1589	0.1732
19	0.0044	0.1408	0.0867	0.1257
20	0.0190	0.0048	0.0429	0.0305
21	0.1850	0.1890	0.1830	0.1149
22	0.0241	0.1251	0.0230	0.1432
23	0.0715	0.0292	0.0954	0.0150
24	0.0986	0.0595	0.1872	0.1989
25	0.0491	0.1936	0.1470	0.1555
26	0.1013	0.1759	0.1406	0.1092
27	0.1586	0.1574	0.0893	0.1800
28	0.0874	0.0131	0.1205	0.0551
29	0.1884	0.0876	0.1726	0.1601
30	0.0298	0.1360	0.0644	0.0791

The curve of LDR' (Figure 9.5 Red solid line) and the curve of LDR (Figure 9.5 Blue solid line) were similar, which support the feasibility of the method using estimated LDR (LDR') to represent LDR.

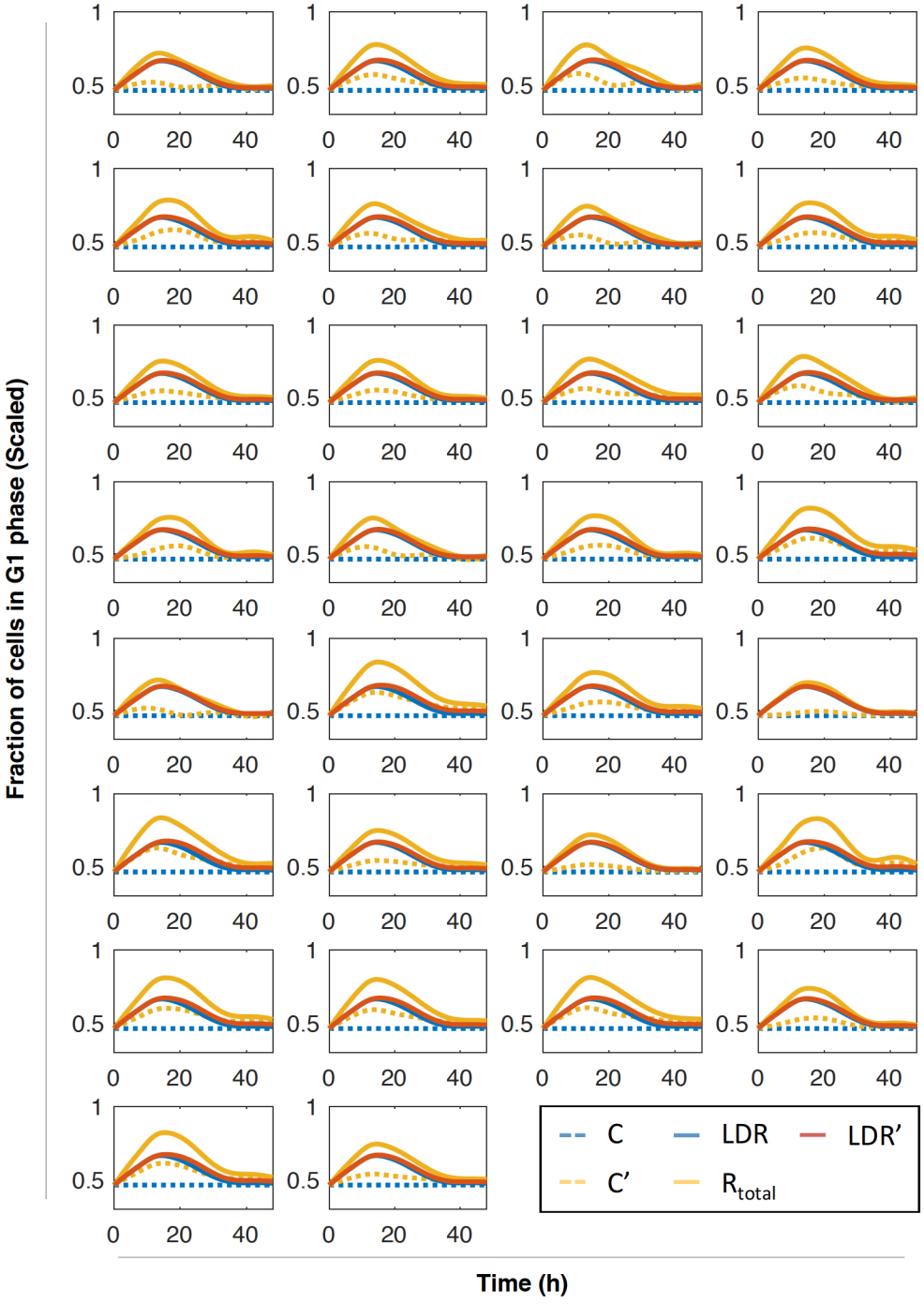


Figure 9.5 Validation of the method of representing LDR (ligand dependent response) by estimated LDR (LDR'). Blue dash line: C, data of control group; Blue solid line: LDR, data of cells upon 50pM TGFβ treatment; Yellow dash line: C', data of control group with LIF (ligand independent fluctuations); Yellow solid line: R_{total}, data of TGFβ-treated group with LIF; Red solid line: LDR', estimated LDR.

List of Figures

List of Tables

Bibliography

1. Hanahan, D. & Weinberg, R.A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674 (2011).
2. Fraguas, S., Barberán, S. & Cebrià, F. EGFR signaling regulates cell proliferation, differentiation and morphogenesis during planarian regeneration and homeostasis. in *Developmental Biology*, Vol. 354 87-101 (2011).
3. Su, N., Jin, M. & Chen, L. Role of FGF/FGFR signaling in skeletal development and homeostasis: learning from mouse models. in *Bone Research*, Vol. 2 1 (2014).
4. Russell, J.O. & Monga, S.P. Wnt/β-Catenin Signaling in Liver Development, Homeostasis, and Pathobiology. in *Annual Review of Pathology: Mechanisms of Disease*, Vol. 13 351-378 (Annual Reviews, 2018).
5. Wu, M.Y. & Hill, C.S. Tgf-beta superfamily signaling in embryonic development and homeostasis. *Dev Cell* **16**, 329-343 (2009).
6. Chizzolini, C., Dayer, J.M. & Miossec, P. Cytokines in chronic rheumatic diseases: is everything lack of homeostatic balance? *Arthritis Res Ther* **11**(2009).
7. Miyajima, A., Ito, Y. & Kinoshita, T. Cytokine signaling for proliferation, survival, and death in hematopoietic cells. *Int J Hematol* **69**, 137-146 (1999).
8. Massagué, J. TGFβ signalling in context. in *Nat. Rev. Mol. Cell Biol.*, Vol. 13 616-630 (2012).
9. Siegel, P.M. & Massagué, J. Cytostatic and apoptotic actions of TGF-β in homeostasis and cancer. in *Nat Rev Cancer*, Vol. 3 807-820 (2003).
10. Massagué, J. TGFβ in Cancer. in *Cell*, Vol. 134 215-230 (2008).
11. Moses, H.L., Roberts, A.B. & Derynck, R. The Discovery and Early Days of TGF-β: A Historical Perspective. in *Cold Spring Harb Perspect Biol*, Vol. 8 a021865 (2016).
12. Shi, Y.G. & Massague, J. Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell* **113**, 685-700 (2003).
13. Hill, C.S. Nucleocytoplasmic shuttling of Smad proteins. *Cell Res* **19**, 36-46 (2009).

14. Varelas, X., *et al.* TAZ controls Smad nucleocytoplasmic shuttling and regulates human embryonic stem-cell self-renewal. *Nat Cell Biol* **10**, 837-848 (2008).

15. Wrighton, K.H., Dai, F. & Feng, X.H. A new kid on the TGFbeta block: TAZ controls Smad nucleocytoplasmic shuttling. *Dev Cell* **15**, 8-10 (2008).

16. Massague, J. & Wotton, D. Transcriptional control by the TGF-beta/Smad signaling system. *EMBO J* **19**, 1745-1754 (2000).

17. Chen, C.R., Kang, Y., Siegel, P.M. & Massague, J. E2F4/5 and p107 as Smad cofactors linking the TGFbeta receptor to c-myc repression. *Cell* **110**, 19-32 (2002).

18. Kang, Y., Chen, C.R. & Massague, J. A self-enabling TGFbeta response coupled to stress signaling: Smad engages stress response factor ATF3 for Id1 repression in epithelial cells. *Mol Cell* **11**, 915-926 (2003).

19. Warmflash, A., *et al.* Dynamics of TGF-β signaling reveal adaptive and pulsatile behaviors reflected in the nuclear localization of transcription factor Smad4. in *Proc. Natl. Acad. Sci. U.S.A.*, Vol. 109 E1947-1956 (2012).

20. Inman, G.J., Nicolas, F.J. & Hill, C.S. Nucleocytoplasmic shuttling of Smads 2, 3, and 4 permits sensing of TGF-beta receptor activity. *Molecular Cell* **10**, 283-294 (2002).

21. Costanza, B., Umelo, I.A., Bellier, J., Castronovo, V. & Turtoi, A. Stromal Modulators of TGF-beta in Cancer. *J Clin Med* **6**(2017).

22. Zhang, Y.E. Non-Smad pathways in TGF-beta signaling. *Cell Res* **19**, 128-139 (2009).

23. Choi, M.E., Ding, Y. & Kim, S.I. TGF-beta signaling via TAK1 pathway: role in kidney fibrosis. *Semin Nephrol* **32**, 244-252 (2012).

24. Gorska, A.E., Joseph, H., Derynck, R., Moses, H.L. & Serra, R. Dominant-negative interference of the transforming growth factor beta type II receptor in mammary gland epithelium results in alveolar hyperplasia and differentiation in virgin mice. *Cell Growth Differ* **9**, 229-238 (1998).

25. Bottinger, E.P., Jakubczak, J.L., Haines, D.C., Bagnall, K. & Wakefield, L.M. Transgenic mice overexpressing a dominant-negative mutant type II transforming growth factor beta receptor show enhanced tumorigenesis in the mammary gland and lung in response to the carcinogen 7,12-dimethylbenz-[alpha]-anthracene. *Cancer Research* **57**, 5564-5570 (1997).

26. Joseph, H., Gorska, A.E., Sohn, P., Moses, H.L. & Serra, R. Overexpression of a kinase-deficient transforming growth factor-beta type II receptor in mouse mammary stroma results in increased epithelial branching. *Molecular Biology of the Cell* **10**, 1221-1234 (1999).

27. Nguyen, A.V. & Pollard, J.W. Transforming growth factor beta 3 induces cell death during the first stage of mammary gland involution. *Development* **127**, 3107-3118 (2000).

28. Cui, W., *et al.* Concerted Action of Tgf-Beta-1 and Its Type-Ii Receptor in Control of Epidermal Homeostasis in Transgenic Mice. *Gene Dev* **9**, 945-955 (1995).

29. Wang, X.J., Liefer, K.M., Tsai, S., O'Malley, B.W. & Roop, D.R. Development of gene-switch transgenic mice that inducibly express transforming growth factor beta 1 in the epidermis. *P Natl Acad Sci USA* **96**, 8483-8488 (1999).

30. Liu, X., *et al.* Conditional epidermal expression of TGF beta 1 blocks neonatal lethality but causes a reversible hyperplasia and alopecia. *P Natl Acad Sci USA* **98**, 9139-9144 (2001).

31. Dickson, M.C., *et al.* Defective haematopoiesis and vasculogenesis in transforming growth factor-beta 1 knock out mice. *Development* **121**, 1845-1854 (1995).

32. Choi, M.E. & Ballermann, B.J. Inhibition of capillary morphogenesis and associated apoptosis by dominant negative mutant transforming growth factor-beta receptors. *J Biol Chem* **270**, 21144-21150 (1995).

33. Goumans, M.J., *et al.* Balancing the activation state of the endothelium via two distinct TGF-beta type I receptors. *EMBO J* **21**, 1743-1753 (2002).

34. Brabletz, T., *et al.* Transforming growth factor beta and cyclosporin A inhibit the inducible activity of the interleukin-2 gene in T cells through a noncanonical octamer-binding site. *Mol Cell Biol* **13**, 1155-1162 (1993).

35. Sillett, H.K., Cruickshank, S.M., Southgate, J. & Trejdosiewicz, L.K. Transforming growth factor-beta promotes 'death by neglect' in post-activated human T cells. *Immunology* **102**, 310-316 (2001).

36. Gorelik, L. & Flavell, R.A. Transforming growth factor-beta in T-cell biology. *Nat Rev Immunol* **2**, 46-53 (2002).

37. Morgan, D.O. Principles of Cdk Regulation. *Nature* **374**, 131-134 (1995).

38. John, P.C.L., Mews, M. & Moore, R. Cyclin/Cdk complexes: their involvement in cell cycle progression and mitotic division. *Protoplasma* **216**, 119-142 (2001).

39. Reed, S.I. Control of the G1/S transition. *Cancer Surv* **29**, 7-23 (1997).

40. Reynisdottir, I. & Massague, J. The subcellular locations of p15(Ink4b) and p27(Kip1) coordinate their inhibitory interactions with cdk4 and cdk2. *Genes Dev* **11**, 492-503 (1997).

41. Harper, J.W., *et al.* Inhibition of cyclin-dependent kinases by p21. *Mol Biol Cell* **6**, 387-400 (1995).

42. Adams, P.D., *et al.* Identification of a cyclin-cdk2 recognition motif present in substrates and p21-like cyclin-dependent kinase inhibitors. *Mol Cell Biol* **16**, 6623-6633 (1996).

43. Xiong, Y., *et al.* P21 Is a Universal Inhibitor of Cyclin Kinases. *Nature* **366**, 701-704 (1993).

44. Herkert, B. & Eilers, M. Transcriptional repression: the dark side of myc. *Genes Cancer* **1**, 580-586 (2010).

45. Wanzel, M., Herold, S. & Eilers, M. Transcriptional repression by Myc. *Trends Cell Biol* **13**, 146-150 (2003).

46. Lasorella, A., Noseda, M., Beyna, M., Yokota, Y. & Iavarone, A. Id2 is a retinoblastoma protein target and mediates signalling by Myc oncoproteins. *Nature* **407**, 592-598 (2000).

47. Elowitz, M.B., Levine, A.J., Siggia, E.D. & Swain, P.S. Stochastic gene expression in a single cell. *Science* **297**, 1183-1186 (2002).

48. Cheong, R., Paliwal, S. & Levchenko, A. Models at the single cell level. *Wires Syst Biol Med* **2**, 34-48 (2010).

49. Spencer, S.L., Gaudet, S., Albeck, J.G., Burke, J.M. & Sorger, P.K. Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature* **459**, 428-432 (2009).

50. Baxter, A.E., Niessl, J., Morou, A. & Kaufmann, D.E. RNA flow cytometric FISH for investigations into HIV immunology, vaccination and cure strategies. *AIDS Res Ther* **14**, 40 (2017).

51. Grau-Exposito, J.*, et al.* A Novel Single-Cell FISH-Flow Assay Identifies Effector Memory CD4(+) T cells as a Major Niche for HIV-1 Transcription in HIV-Infected Patients. *MBio* **8**(2017).

52. Arrigucci, R.*, et al.* FISH-Flow, a protocol for the concurrent detection of mRNA and protein in single cells using fluorescence in situ hybridization and flow cytometry. *Nat Protoc* **12**, 1245-1260 (2017).

53. Zahedipour, F.*, et al.* Development of Flow Cytometry-Fluorescent In Situ Hybridization (Flow-FISH) Method for Detection of PML/RARa Chromosomal Translocation in Acute Promyelocytic Leukemia Cell Line. *Avicenna J Med Biotechnol* **9**, 104-108 (2017).

54. Hu, P., Zhang, W., Xin, H. & Deng, G. Single Cell Isolation and Analysis. *Front Cell Dev Biol* **4**, 116 (2016).

55. Muzzey, D. & van Oudenaarden, A. Quantitative time-lapse fluorescence microscopy in single cells. *Annu Rev Cell Dev Biol* **25**, 301-327 (2009).

56. Reynisdottir, I., Polyak, K., Iavarone, A. & Massague, J. Kip/Cip and Ink4 Cdk inhibitors cooperate to induce cell cycle arrest in response to TGF-beta. *Genes Dev* **9**, 1831-1845 (1995).

57. Mukherjee, P., Winter, S.L. & Alexandrow, M.G. Cell cycle arrest by transforming growth factor beta1 near G1/S is mediated by acute abrogation of prereplication complex activation involving an Rb-MCM interaction. *Mol Cell Biol* **30**, 845-856 (2010).

58. Shipley, G.D., Pittelkow, M.R., Wille, J.J., Jr., Scott, R.E. & Moses, H.L. Reversible inhibition of normal human prokeratinocyte proliferation by type beta transforming growth factor-growth inhibitor in serum-free medium. *Cancer Res* **46**, 2068-2071 (1986).

59. Heimark, R.L., Twardzik, D.R. & Schwartz, S.M. Inhibition of endothelial regeneration by type-beta transforming growth factor from platelets. *Science* **233**, 1078-1080 (1986).

60. Lee, J., Choi, J.H. & Joo, C.K. TGF-beta1 regulates cell fate during epithelial-mesenchymal transition by upregulating survivin. *Cell Death Dis* **4**, e714 (2013).

61. Laiho, M., DeCaprio, J.A., Ludlow, J.W., Livingston, D.M. & Massague, J. Growth inhibition by TGF-beta linked to suppression of retinoblastoma protein phosphorylation. *Cell* **62**, 175-185 (1990).

62. Pardee, A.B. & Keyomarsi, K. Modification of cell proliferation with inhibitors. *Curr Opin Cell Biol* **4**, 186-191 (1992).

63. Darzynkiewicz, Z., Halicka, H.D., Zhao, H. & Podhorecka, M. Cell synchronization by inhibitors of DNA replication induces replication stress and DNA damage response: analysis by flow cytometry. *Methods Mol Biol* **761**, 85-96 (2011).

64. Cheong, R., Rhee, A., Wang, C.J., Nemenman, I. & Levchenko, A. Information Transduction Capacity of Noisy Biochemical Signaling Networks. *Science* **334**, 354-358 (2011).

65. Ladbury, J.E. & Arold, S.T. Noise in cellular signaling pathways: causes and effects. *Trends Biochem Sci* **37**, 173-178 (2012).

66. Selimkhanov, J.*, et al.* Systems biology. Accurate information transmission through dynamic biochemical signaling networks. *Science* **346**, 1370-1373 (2014).

67. Sonnen, K.F. & Aulehla, A. Dynamic signal encoding--from cells to organisms. *Semin Cell Dev Biol* **34**, 91-98 (2014).

68. Spiller, D.G., Wood, C.D., Rand, D.A. & White, M.R. Measurement of single-cell dynamics. *Nature* **465**, 736-745 (2010).

69. Purvis, J.E. & Lahav, G. Encoding and decoding cellular information through signaling dynamics. *Cell* **152**, 945-956 (2013).

70. Purvis, J.E. & Lahav, G. Encoding and Decoding Cellular Information through Signaling Dynamics. in *Cell*, Vol. 152 945-956 (Elsevier Inc., 2013).

71. Hamill, O.P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F.J. Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches. *Pflugers Arch* **391**, 85-100 (1981).

72. Fertig, N., Blick, R.H. & Behrends, J.C. Whole cell patch clamp recording performed on a planar glass chip. *Biophys J* **82**, 3056-3062 (2002).

73. Wijgerde, M., Grosveld, F. & Fraser, P. Transcription complex stability and chromatin dynamics in vivo. *Nature* **377**, 209-213 (1995).

74. Freudiger, C.W.*, et al.* Label-free biomedical imaging with high sensitivity by stimulated Raman scattering microscopy. *Science* **322**, 1857-1861 (2008).

75. Evans, C.L. & Xie, X.S. Coherent anti-stokes Raman scattering microscopy: chemical imaging for biology and medicine. *Annu Rev Anal Chem (Palo Alto Calif)* **1**, 883-909 (2008).

76. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W. & Prasher, D.C. Green Fluorescent Protein as a Marker for Gene-Expression. *Science* **263**, 802-805 (1994).

77. Day, R.N. & Davidson, M.W. The fluorescent protein palette: tools for cellular imaging (vol 38, pg 2887, 2009). *Chem Soc Rev* **40**, 5923-5923 (2011).

78. Carrero, G., McDonald, D., Crawford, E., de Vries, G. & Hendzel, M.J. Using FRAP and mathematical modeling to determine the in vivo kinetics of nuclear proteins. *Methods* **29**, 14-28 (2003).

79. Day, R.N. & Davidson, M.W. The fluorescent protein palette: tools for cellular imaging. *Chem Soc Rev* **38**, 2887-2921 (2009).

80. Kim, S.A., Heinze, K.G. & Schwille, P. Fluorescence correlation spectroscopy in living cells. *Nat Methods* **4**, 963-973 (2007).

81. Brown, C.M., Wiseman, P.W., Webb, D.J., Hebert, B. & Horwitz, A.R. Spatial mapping of integrin, paxillin and alpha-actinin interactions and dynamics during cell migration by image correlation microscopy. *Molecular Biology of the Cell* **15**, 161a-161a (2004).

82. Gu, Y., Di, W.L., Kelsell, D.P. & Zicha, D. Quantitative fluorescence resonance energy transfer (FRET) measurement with acceptor photobleaching and spectral unmixing. *J Microsc* **215**, 162-173 (2004).

83. Suhling, K., French, P.M. & Phillips, D. Time-resolved fluorescence microscopy. *Photochem Photobiol Sci* **4**, 13-22 (2005).

84. Michnick, S.W., Ear, P.H., Manderson, E.N., Remy, I. & Stefan, E. Universal strategies in research and drug discovery based on protein-fragment complementation assays. *Nat Rev Drug Discov* **6**, 569-582 (2007).

85. Ozawa, T., Natori, Y., Sato, M. & Umezawa, Y. Imaging dynamics of endogenous mitochondrial RNA in single living cells. *Nat Methods* **4**, 413-419 (2007).

86. Hida, N., *et al.* High-Sensitivity Real-Time Imaging of Dual Protein-Protein Interactions in Living Subjects Using Multicolor Luciferases. *Plos One* **4**(2009).

87. Batchelor, E., Loewer, A., Mock, C. & Lahav, G. Stimulus-dependent dynamics of p53 in single cells. *Mol Syst Biol* **7**, 488 (2011).

88. Batchelor, E., Mock, C.S., Bhan, I., Loewer, A. & Lahav, G. Recurrent initiation: a mechanism for triggering p53 pulses in response to DNA damage. *Mol Cell* **30**, 277-289 (2008).

89. Nelson, D.E., *et al.* Oscillations in NF-kappaB signaling control the dynamics of gene expression. *Science* **306**, 704-708 (2004).

90. Strasen, J., *et al.* Cell-specific responses to the cytokine TGF beta are determined by variability in protein levels. *Molecular Systems Biology* **14**(2018).

91. Frick, C.L., Yarka, C., Nunns, H. & Goentoro, L. Sensing relative signal in the Tgf-beta/Smad pathway. *P Natl Acad Sci USA* **114**, E2975-E2982 (2017).

92. Oki, T., *et al.* A novel cell-cycle-indicator, mVenus-p27K-, identifies quiescent cells and visualizes G0-G1 transition. *Sci Rep* **4**, 4012 (2014).

93. Barr, A.R., Heldt, F.S., Zhang, T., Bakal, C. & Novak, B. A Dynamical Framework for the All-or-None G1/S Transition. *Cell Syst* **2**, 27-37 (2016).

94. Wilson, K.A., Elefanty, A.G., Stanley, E.G. & Gilbert, D.M. Spatio-temporal re-organization of replication foci accompanies replication domain consolidation during human pluripotent stem cell lineage specification. *Cell Cycle* **15**, 2464-2475 (2016).

95. Barr, A.R., *et al.* DNA damage during S-phase mediates the proliferation-quiescence decision in the subsequent G1 via p21 expression. *Nat Commun* **8**, 14728 (2017).

96. Chao, H.X., *et al.* Orchestration of DNA Damage Checkpoint Dynamics across the Human Cell Cycle. *Cell Syst* **5**, 445-459 e445 (2017).

97. Theobald, B., *et al.* Suppression of Ser/Thr phosphatase 4 (PP4C/PPP4C) mimics a novel post-mitotic action of fostriecin, producing mitotic slippage followed by tetraploid cell death. *Mol Cancer Res* **11**, 845-855 (2013).

98. Sakaue-Sawano, A., *et al.* Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. *Cell* **132**, 487-498 (2008).

99. Bajar, B.T., *et al.* Fluorescent indicators for simultaneous reporting of all four cell cycle phases. *Nat Methods* **13**, 993-996 (2016).

100. Zielke, N. & Edgar, B.A. FUCCI sensors: powerful new tools for analysis of cell proliferation. *Wiley Interdiscip Rev Dev Biol* **4**, 469-487 (2015).

101. Newman, R.H. & Zhang, J. Fucci: street lights on the road to mitosis. *Chem Biol* **15**, 97-98 (2008).

102. Nishitani, H., Lygerou, Z., Nishimoto, T. & Nurse, P. The Cdt1 protein is required to license DNA for replication in fission yeast. *Nature* **404**, 625-628 (2000).

103. Sakaue-Sawano, A., *et al.* Visualizing Spatiotemporal Dynamics of Multicellular Cell-Cycle Progression. in *Cell*, Vol. 132 487-498 (2008).

104. Anderson, A.R. & Quaranta, V. Integrative mathematical oncology. *Nat Rev Cancer* **8**, 227-234 (2008).

105. Weis, M.C., Avva, J., Jacobberger, J.W. & Sreenath, S.N. A data-driven, mathematical model of mammalian cell cycle regulation. *PLoS One* **9**, e97130 (2014).

106. Singhania, R., Sramkoski, R.M., Jacobberger, J.W. & Tyson, J.J. A hybrid model of mammalian cell cycle regulation. *PLoS Comput Biol* **7**, e1001077 (2011).

107. Sherer, E., Hannemann, R.E., Rundell, A. & Ramkrishna, D. Analysis of resonance chemotherapy in leukemia treatment via multi-staged population balance models. *J Theor Biol* **240**, 648-661 (2006).

108. Daukste, L., Basse, B., Baguley, B.C. & Wall, D.J. Mathematical determination of cell population doubling times for multiple cell lines. *Bull Math Biol* **74**, 2510-2534 (2012).

109. Florian, J.A., Jr., Eiseman, J.L. & Parker, R.S. Accounting for quiescent cells in tumour growth and cancer treatment. *Syst Biol (Stevenage)* **152**, 185-192 (2005).

110. Garcia Munzer, D.G., Kostoglou, M., Georgiadis, M.C., Pistikopoulos, E.N. & Mantalaris, A. Cyclin and DNA distributed cell cycle model for GS-NS0 cells. *PLoS Comput Biol* **11**, e1004062 (2015).

111. Nurse, P. A long twentieth century of the cell cycle and beyond. *Cell* **100**, 71-78 (2000).

112. Zhou, B.B. & Elledge, S.J. The DNA damage response: putting checkpoints in perspective. *Nature* **408**, 433-439 (2000).

113. Musacchio, A. & Hardwick, K.G. The spindle checkpoint: structural insights into dynamic signalling. *Nat Rev Mol Cell Biol* **3**, 731-741 (2002).

114. Hartwell, L.H. & Weinert, T.A. Checkpoints: controls that ensure the order of cell cycle events. *Science* **246**, 629-634 (1989).

115. Murray, A.W. Creative blocks: cell-cycle checkpoints and feedback controls. *Nature* **359**, 599-604 (1992).

116. Novak, B. & Tyson, J.J. Numerical analysis of a comprehensive model of M-phase control in Xenopus oocyte extracts and intact embryos. *J Cell Sci* **106 (Pt 4)**, 1153-1168 (1993).

117. Pomerening, J.R., Sontag, E.D. & Ferrell, J.E., Jr. Building a cell cycle oscillator: hysteresis and bistability in the activation of Cdc2. *Nat Cell Biol* **5**, 346-351 (2003).

118. Sha, W., *et al.* Hysteresis drives cell-cycle transitions in Xenopus laevis egg extracts. *Proc Natl Acad Sci U S A* **100**, 975-980 (2003).

119. Gerard, C. & Goldbeter, A. Temporal self-organization of the cyclin/Cdk network driving the mammalian cell cycle. *Proc Natl Acad Sci U S A* **106**, 21643-21648 (2009).

120. Gerard, C. & Goldbeter, A. From quiescence to proliferation: Cdk oscillations drive the mammalian cell cycle. *Front Physiol* **3**, 413 (2012).

121. Qu, Z., Weiss, J.N. & MacLellan, W.R. Regulation of the mammalian cell cycle: a model of the G1-to-S transition. *Am J Physiol Cell Physiol* **284**, C349-364 (2003).

122. Novak, B. & Tyson, J.J. A model for restriction point control of the mammalian cell cycle. *J Theor Biol* **230**, 563-579 (2004).

123. Swat, M., Kel, A. & Herzel, H. Bifurcation analysis of the regulatory modules of the mammalian G1/S transition. *Bioinformatics* **20**, 1506-1511 (2004).

124. Pfeuty, B. Strategic cell-cycle regulatory features that provide mammalian cells with tunable G1 length and reversible G1 arrest. *PLoS One* **7**, e35291 (2012).

125. Aguda, B.D. A quantitative analysis of the kinetics of the G(2) DNA damage checkpoint system. *Proc Natl Acad Sci U S A* **96**, 11352-11357 (1999).

126. Saitou, T. & Imamura, T. Quantitative imaging with Fucci and mathematics to uncover temporal dynamics of cell cycle progression. *Dev Growth Differ* **58**, 6-15 (2016).

127. Staudte, R.G., Guiguet, M. & d'Hooghe, M.C. Additive models for dependent cell populations. *J Theor Biol* **109**, 127-146 (1984).

128. Cowan, R. & Staudte, R. The bifurcating autoregression model in cell lineage studies. *Biometrics* **42**, 769-783 (1986).

129. Smith, J.A. & Martin, L. Do cells cycle? *Proc Natl Acad Sci U S A* **70**, 1263-1267 (1973).

130. Shields, R. Transition probability and the origin of variation in the cell cycle. *Nature* **267**, 704-707 (1977).

131. Shields, R. Further evidence for a random transition in the cell cycle. *Nature* **273**, 755-758 (1978).

132. Shields, R. & Smith, J.A. Cells regulate their proliferation through alterations in transition probability. *J Cell Physiol* **91**, 345-355 (1977).

133. Altinok, A., Gonze, D., Levi, F. & Goldbeter, A. An automaton model for the cell cycle. in *Interface Focus*, Vol. 1 36-47 (2010).

134. Fuentes-Garí, M., *et al.* Selecting a Differential Equation Cell Cycle Model for Simulating Leukemia Treatment. in *Industrial & Engineering Chemistry Research*, Vol. 54 8847-8859 (American Chemical Society, 2015).

135. Fuentes-Gari, M., *et al.* A mathematical model of subpopulation kinetics for the deconvolution of leukaemia heterogeneity. *J R Soc Interface* **12**, 20150276 (2015).

136. Altinok, A., Gonze, D., Levi, F. & Goldbeter, A. An automaton model for the cell cycle. *Interface Focus* **1**, 36-47 (2011).

137. Okumura, Y., Onozawa, M., Morita, T. & Matsuzawa, T. A simple method for estimation of cell cycle parameters. *Exp Cell Res* **78**, 233-236 (1973).

138. Barrett, J.C. A mathematical model of the mitotic cycle and its application to the interpretation of percentage labeled mitoses data. *J Natl Cancer Inst* **37**, 443-450 (1966).

139. Peña-Diaz, J.*, et al.* Transcription profiling during the cell cycle shows that a subset of Polycomb-targeted genes is upregulated during DNA replication. in *Nucleic Acids Res.*, Vol. 41 2846-2856 (2013).

140. Altinok, A., Lévi, F. & Goldbeter, A. A cell cycle automaton model for probing circadian patterns of anticancer drug delivery. in *Adv. Drug Deliv. Rev.*, Vol. 59 1036-1053 (2007).

141. Raue, A.*, et al.* Data2Dynamics: a modeling environment tailored to parameter estimation in dynamical systems. *Bioinformatics* **31**, 3558-3560 (2015).

142. Maiwald, T., Eberhardt, O. & Blumberg, J. Mathematical modeling of biochemical systems with PottersWheel. *Methods Mol Biol* **880**, 119-138 (2012).

143. Iman, R.L. Latin Hypercube Sampling. (John Wiley & Sons, Inc., 2004).

144. Ma, W., Trusina, A., El-Samad, H., Lim, W.A. & Tang, C. Defining network topologies that can achieve biochemical adaptation. in *Cell*, Vol. 138 760-773 (2009).

145. Mckay, M.D., Beckman, R.J. & Conover, W.J. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics* **21**, 239-245 (1979).

146. McQuin, C.*, et al.* CellProfiler 3.0: Next-generation image processing for biology. *PLoS Biol* **16**, e2005970 (2018).

147. Carpenter, A.E.*, et al.* CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol* **7**, R100 (2006).

148. Krenning, L., Feringa, F.M., Shaltiel, I.A., van den Berg, J. & Medema, R.H. Transient activation of p53 in G2 phase is sufficient to induce senescence. in *Molecular Cell*, Vol. 55 59-72 (2014).

149. Zielke, N., Edgar, B.A. & DePamphilis, M.L. Endoreplication. in *Cold Spring Harb Perspect Biol*, Vol. 5 a012948-a012948 (Cold Spring Harbor Lab, 2013).

150. Altschuler, S.J. & Wu, L.F. Cellular Heterogeneity: Do Differences Make a Difference? in *Cell*, Vol. 141 559-563 (2010).

151. Losick, R. & Desplan, C. Stochasticity and cell fate. *Science* **320**, 65-68 (2008).

152. Balaban, N.Q., Merrin, J., Chait, R., Kowalik, L. & Leibler, S. Bacterial persistence as a phenotypic switch. *Science* **305**, 1622-1625 (2004).

153. van Niekerk, G., Davids, L.M., Hattingh, S.M. & Engelbrecht, A.M. Cancer stem cells: A product of clonal evolution? *Int J Cancer* **140**, 993-999 (2017).

154. Tay, S.*, et al.* Single-cell NF-kappaB dynamics reveal digital activation and analogue information processing. *Nature* **466**, 267-271 (2010).

155. Wertek, F. & Xu, C. Digital response in T cells: to be or not to be. *Cell Res* **24**, 265-266 (2014).

156. Kingeter, L.M., Paul, S., Maynard, S.K., Cartwright, N.G. & Schaefer, B.C. Cutting edge: TCR ligation triggers digital activation of NF-kappaB. *J Immunol* **185**, 4520-4524 (2010).

157. Podtschaske, M.*, et al.* Digital NFATc2 activation per cell transforms graded T cell receptor activation into an all-or-none IL-2 expression. *PLoS One* **2**, e935 (2007).

158. Das, J.*, et al.* Digital signaling and hysteresis characterize ras activation in lymphoid cells. *Cell* **136**, 337-351 (2009).

159. Altan-Bonnet, G. & Germain, R.N. Modeling T cell antigen discrimination based on feedback control of digital ERK responses. *PLoS Biol* **3**, e356 (2005).

160. Clark, B. & Hausser, M. Neural coding: hybrid analog and digital signalling in axons. *Curr Biol* **16**, R585-588 (2006).

161. Monod, J. & Jacob, F. Teleonomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harb Symp Quant Biol* **26**, 389-401 (1961).

162. Ferrell, J.E., Jr. Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Curr Opin Cell Biol* **14**, 140-148 (2002).

163. Reyes, J.*, et al.* Fluctuations in p53 Signaling Allow Escape from Cell-Cycle Arrest. in *Molecular Cell* (2018).

164. Coleman, K.E.*, et al.* Sequential replication-coupled destruction at G1/S ensures genome stability. *Gene Dev* **29**, 1734-1746 (2015).

165. Spencer, S.L.*, et al.* The proliferation-quiescence decision is controlled by a bifurcation in CDK2 activity at mitotic exit. in *Cell*, Vol. 155 369-383 (2013).

166. Overton, K.W., Spencer, S.L., Noderer, W.L., Meyer, T. & Wang, C.L. Basal p21 controls population heterogeneity in cycling and quiescent cell cycle states. in *Proc. Natl. Acad. Sci. U.S.A.*, Vol. 111 E4386-4393 (National Academy of Sciences, 2014).

167. Chen, Y.G. Endocytic regulation of TGF-beta signaling. *Cell Res* **19**, 58-70 (2009).

168. Centrella, M., Ji, C., Casinghino, S. & McCarthy, T.L. Rapid flux in transforming growth factor-beta receptors on bone cells. *J Biol Chem* **271**, 18616-18622 (1996).

169. Koli, K.M. & Arteaga, C.L. Processing of the transforming growth factor beta type I and II receptors. Biosynthesis and ligand-induced regulation. *J Biol Chem* **272**, 6423-6427 (1997).

170. Wells, R.G., Yankelev, H., Lin, H.Y. & Lodish, H.F. Biosynthesis of the type I and type II TGF-beta receptors. Implications for complex formation. *J Biol Chem* **272**, 11444-11451 (1997).

171. Zhang, L.*, et al.* Zebrafish Dpr2 inhibits mesoderm induction by promoting degradation of nodal receptors. *Science* **306**, 114-117 (2004).

172. Mitchell, H., Choudhury, A., Pagano, R.E. & Leof, E.B. Ligand-dependent and -independent transforming growth factor-beta receptor recycling regulated by clathrin-mediated endocytosis and Rab11. *Mol Biol Cell* **15**, 4166-4178 (2004).

173. Di Guglielmo, G.M., Le Roy, C., Goodfellow, A.F. & Wrana, J.L. Distinct endocytic pathways regulate TGF-beta receptor signalling and turnover. *Nat Cell Biol* **5**, 410-421 (2003).

174. Vizan, P.*, et al.* Controlling long-term signaling: receptor dynamics determine attenuation and refractory behavior of the TGF-beta pathway. *Sci Signal* **6**, ra106 (2013).

175. Datto, M.B.*, et al.* Transforming growth factor beta induces the cyclin-dependent kinase inhibitor p21 through a p53-independent mechanism. *Proc Natl Acad Sci U S A* **92**, 5545-5549 (1995).

176. Pellegata, N.S., Antoniono, R.J., Redpath, J.L. & Stanbridge, E.J. DNA damage and p53-mediated cell cycle arrest: A reevaluation. *P Natl Acad Sci USA* **93**, 15209-15214 (1996).

177. Chao, H.X.*, et al.* Orchestration of DNA Damage Checkpoint Dynamics across the Human Cell Cycle. in *Cell Syst*, Vol. 5 445-459.e445 (2017).

178. Schindelin, J.*, et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-682 (2012).

179. Warmflash, A.*, et al.* Dynamics of TGF-beta signaling reveal adaptive and pulsatile behaviors reflected in the nuclear localization of transcription factor Smad4. *Proc Natl Acad Sci U S A* **109**, E1947-1956 (2012).

Acknowledgements

First of all, I would like to thank my PhD supervisor, Dr. Zhike Zi for giving me the opportunity to work in his lab in Max Planck Institute for Molecular Genetics and to join the research training group of Computational Systems Biology (CSB). I learnt a lot during the three years of study.

I also want to thank Prof. Dr. Dr. h.c. Edda Klipp, who is very supportive and caring. Thank you for participating my Thesis Advisory Committee meetings, reviewing the project progress reports and providing suggestions during my doctoral study. I would like to thank Dr. Jana Wolf for reviewing this thesis and providing feedbacks.

Next, many thanks belong to my colleagues in the lab, especially to Yuchao who gave me a lot of supports in research and joined me on many adventures in daily life, and to Susanne for her assistance in cell culture. I also would like to thank Marjin, Difan, Hongqing and Dan who provided a friendly working environment. I also appreciate the technical assistance from the Microscopy and IT facility in the Max Planck Institute for Molecular Genetics.

Special thanks belong to CSB, especially to Cordelia and Marylu for all the help and caring. Thanks for providing the opportunities to meet many researchers around the world. I would like to say thanks to all the members for the relaxed and open atmosphere on science.

Last but not least, I would like to thank my dear family and wonderful friends. Especially, I want to thank my parents for the endless love, support and encouragement throughout my life.

Statement

Ich erkläre, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Hilfsmittel angefertigt habe.

Guoyu Wu

Berlin, 10.10.2018